QUANTITATIVE DETECTION AND DELINEATION OF HEAD AND NECK CANCER USING HYPERSPECTRAL IMAGING AND MACHINE LEARNING

A Dissertation
Presented to
The Academic Faculty

by

Guolan Lu

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Biomedical Engineering

Georgia Institute of Technology
December, 2016

QUANTITATIVE DETECTION AND DELINEATION OF HEAD AND NECK CANCER USING HYPERSPECTRAL IMAGING AND MACHINE LEARNING

Approved by:

Dr. Baowei Fei, Advisor
Wallace H. Coulter Department of Biomedical Engineering
Georgia Institute of Technology and Emory University
Department of Radiology and Imaging Sciences, School of Medicine
Emory University

Dr. Georgia Z. Chen
School of Medicine
Emory University

Dr. John N. Oshinski
Wallace H. Coulter Department of Biomedical Engineering
Georgia Institute of Technology and Emory University

Dr. Brian W. Pogue
Thayer School of Engineering
Dartmouth College

Dr. May D. Wang
Wallace H. Coulter Department of Biomedical Engineering
Georgia Institute of Technology and Emory University

Date Approved: October 17, 2016

To my loving family

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deep appreciation to my advisor Dr. Baowei Fei for the opportunity to work in his lab. His support and guidance for my research and career development, along with great patience, enthusiasm, and immerse knowledge were essential to my success throughout my PhD studies.

I would like to thank my thesis committee members: Dr. Georgia Chen, Dr. John Osinksi, Dr. Brian Pogue, and Dr. May Wang, for their guidance and valuable feedback, and for taking time out of their busy schedule to examine and critique my research.

I would like to thank all my collaborators: Dr. Dongsheng Wang, Dr. Susan Muller, Dr. Amy Chen, Dr. James Little, Dr. Kelly Magliocca, Dr. Hongzheng Zhang, Dr. Xu Wang, all the surgeons and nurses from the department of Otolaryngology, all the pathologist assistants of Emory Midtown hospital, and Ms. Jennifer Shelton from Winship Pathology Core Service, for all their time and help with my thesis project.

I would like to thank all the lab members of the Quantitative BioImaging Laboratory (QBIL):Ms. Luma Halig, Dr. Xulei Qin, Mr. James Dormer, Dr. Zhiqiang Tian, Dr. Lizhi Liu, Ms. Ling Ma, Dr. Yun Zhang, Dr. Rongrong Guo, Dr. Guoyi Zhang, Mr. Martin Halicek, and Dr. Shana Watson for all their help with my research and for their friendship in making my PhD life more colorful and joyous. In addition, I would like to thank all the faculty and staff members in the Center for Systems Imaging (CSI) at Emory University, for all their help and support during my work in CSI.

Finally, I would like to thank my family for everything. I would like to especially thank my parents for the unconditional love, support, and trust, and for providing me the

peace and happiness in this journey. I would like to especially thank my brother Guojun Lv for always being there for me, inspiring me with positive energy, being my biggest advocate, and for helping me through all the difficult times as I was growing up. I would like to thank my husband, Xulei, for his love, patience, encouragement, and continual support of my academic endeavors, and for bringing me so much happiness over the past several years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

ANN	K-nearest neighbor
CV	Cross validation
4NQO	4-Nitroquinoline N-oxide
H&E	Haematoxylin and eosin
HNSCC	Head and neck squamous cell carcinoma
HSI	Hyperspectral imaging
KNN	Artificial neural network
LCTF	Liquid crystal tunable filter
LDA	Linear discriminant analysis
NIR	Near-infrared
PBS	Phosphate-buffered saline
PCA	Principal Component analysis
SVM	Support vector machine
3D	Three-dimensional
2D	Two-dimensional
2-NBDG	2-deoxy-2-[(7-nitro-2,1,3-benzoxadiazol-4-yl)amino]-D-glucose
UV	Ultraviolet
VIS	Visible

SUMMARY

Worldwide, over 500,000 patients are diagnosed with head and neck squamous cell carcinoma each year, posing a substantial economic burden to society [1] [2]. Despite significant advances in cancer treatment, early cancer detection and subsequent surgical resection of tumor remains one of the most promising approaches to improve the survival and quality of life of patients.

As an emerging optical modality, hyperspectral imaging (HSI) holds potential promise for early cancer detection and image-guided surgery. The major advantage of HSI is that it is a noninvasive technology that does not require any contrast agent, and that it combines wide-field imaging and spectroscopy to simultaneously attain both spatial and spectral information from an object in a non-contact way. Light delivered to the tissue surface undergoes multiple elastic scattering and absorption interactions, and part of it returns as diffuse reflectance carrying diagnostic information about the underlying tissue structure and composition. The biochemical and morphological properties of the tissue change during disease progression, which would alter the tissue optical properties. Hyperspectral images contain high-dimensional spectral information at each image point and can be analyzed for visualization, characterization, and quantification of the disease state in biological tissue.

Hyperspectral images contain rich spectral-spatial data, but the ability to pull out important diagnostic information from the large amount of data is a key to fully explore the potential of HSI. Past spectroscopic analysis methods mainly focus on modeling the light propagation in biological tissue with analytical approximation such as the diffusion equation, Monte Carlo model, and empirical methods, and extracting information about tissue structure and composition, such as hemoglobin concentration, hemoglobin oxygen saturation, effective scatterer density, and effective scatterer size. The analytical

modeling of diffuse reflectance links the optical properties with tissue pathophysiology, which provides a biological interpretation for the diagnostic information contained in hyperspectral data.

To fully utilize the spatiospectral information in the hyperspectral data, this dissertation proposed to investigate the use of machine learning methods as quantification tools for predictive analysis of hyperspectral images towards quantitative cancer detection and diagnosis. Machine learning methods, which make no assumptions about biological tissue, can be employed to extract the discriminative spectral-spatial features, learn the underlying patterns, build computational models from large volume of data, and make predictions of cancer in biological tissue. The general procedures of the machine learning methods involve feature extraction and selection, classification, and performance evaluation, providing versatile frameworks for quantitative analysis of hyperspectral images. Incorporating advantages of machine learning technique into the HSI system could provide an objective and sensitive technique suitable for cancer detection.

The overall goal of this dissertation was to investigate the potential of label-free HSI technology combined with machine learning methods as a noninvasive diagnostic tool for quantitative detection and delineation of head and neck cancer. More specifically, this dissertation work has two applications: early detection of cancer, and surgical guidance. To achieve this, we had four different aims. The first two aims evaluated the diagnostic performance of HSI and machine learning algorithms for differentiating cancer from normal tissue in preclinical animal models, including a subcutaneous cancer model (Aim 1), and a chemically-induced tongue carcinogenesis model (Aim 2). The last two aims investigated the detection and delineation of head and neck cancer in a surgical animal model (Aim 3) and fresh surgical specimens of human patients (Aim 4).

First, a spectral-spatial classification method was developed to evaluate the diagnostic performance of HSI as a noninvasive tool to discriminate head and neck cancer from its surrounding healthy tissue *in vivo* in a xenograft cancer model. An

experiment was designed and conducted for acquiring hyperspectral images of subcutaneous head and neck cancer in mice. The quantitative analysis of the hyperspectral data was challenging due to the large data volume, high dimensionality of spectral bands, and high spatial resolution. Therefore, this study proposed a tensor-based spectral-spatial classification method that utilized the entire spectra at each pixel as well as the spectral information from its neighborhood to differentiate cancer from normal tissue. This study demonstrated the feasibility of utilizing HSI for predicting cancerous probability of *in vivo* tissue in an animal model in a noninvasive and non-contact manner.

To further evaluate the potential of HSI as a noninvasive diagnostic tool, we proceeded to a chemically-induced tongue carcinogenesis model in Aim 2. In comparison to the relatively simple cancer model in the first animal study, where the xenograft tumors were visible to the naked eye, the chemically-induced carcinogenesis model had two advantages: (1) it simulated the biology of human tongue cancer development and progression, which provided an opportunity to study the stages of the carcinogenesis process; (2) the induced tongue tumors were barely visible to the naked eye during the progressive development. In this animal study, imaging was conducted to scan both *in vivo* and *ex vivo* dorsal surface of mouse tongues at multiple time points during the carcinogenesis process, and a color-coded pathology map were reconstructed from a series of sagittally-sectioned histological slices of the tongue as the gold standard. Various machine learning classifiers, such as discriminant analysis, ensemble learning methods, and support vector machines, were implemented and validated for hyperspectral images to differentiate non-neoplastic tissue from tongue neoplasia including dysplasia, carcinoma in situ, and carcinoma. Moreover, the diagnostic performance of HSI, autofluorescence imaging, and fluorescence imaging were evaluated and compared in mouse data. This study demonstrated the capability of HSI and machine learning techniques to noninvasively detect and delineate pre-cancerous and cancerous tissue in a preclinical murine carcinogenesis model.

To assess the capability of HSI for head and neck cancer detection in surgical settings, a framework for hyperspectral image processing and quantification was proposed, which included a set of steps involving image preprocessing, glare removal, feature extraction, and image classification. The framework was tested on images from mice with head and neck cancer. The image analysis methods explored the utility of multiple spectral features, including Fourier coefficients, normalized reflectance, mean, and spectral derivatives. The experimental results demonstrated the feasibility of the hyperspectral image processing and quantification framework for cancer detection during the animal tumor surgery.

Finally, a proof-of-principle study was conducted to determine the feasibility of HSI for differentiating tumor from normal tissue in human patients with head and neck cancer. Intraoperative frozen section diagnosis is a widely used practice during head and neck cancer surgery to assess the extent of cancer resection, which is costly in terms of time and personnel and may suffer from sampling errors. Therefore, fresh surgical specimens from a variety of anatomic sites, including oral cavity, larynx, thyroid, parotid, paranasal sinus, and nasal cavity, were collected and imaged. Two machine learning-based frameworks, including intra-patient and inter-patient classification, were developed for the quantitative detection and delineation of head and neck cancer in hyperspectral images, which were validated by the pathological gold standard. First, intra-patient classification was conducted, which built training models from tumor and normal tissue samples and evaluated the predictive performance of the model on the tumor-normal interface tissue samples of the same patient. To account for inter-patient heterogeneity, inter-patient classification was further conducted, which built training models and

evaluated the model performance using tissue samples from different patients. Both intra- and inter- patient classification methods could be translated into the clinic for predicting the margin status of the resected fresh surgical specimen. This proof-of-concept study demonstrated that HSI is capable of detecting and delineating tumors in fresh surgical specimens. Future development of this approach for intraoperative surgical guidance and validation could facilitate the clinical translation of HSI for quantitative surgical margin assessment during surgery.

CHAPTER 1

INTRODUCTION

1.1 Background and Motivation

Cancer is among the leading causes of morbidity and mortality worldwide, with approximately 14.1 million new cancer cases and 8.2 million deaths in 2012 [3]. Head and neck cancer is the sixth most common cancer worldwide. This disease can affect the oral cavity, pharynx, larynx, sinuses, thyroid gland, and salivary gland [4]. Most people who develop head and neck cancer have advanced disease at the time of diagnosis. In the United States, 65% of the patients with oral cavity and pharynx cancer have regional or distant spread of their disease at the time of diagnosis [5]. Moreover, the 5-year survival rate for localized oral cavity and pharynx cancer is 83.0%, but the survival rate drops to only 37.7% for metastasized cancer [5]. The early detection and subsequent surgical resection of tumors represents one of the most promising approaches to reducing the growing cancer burden. The following sections provide the background and motivation of this dissertation.

1.1.1 Early cancer detection

Cancer is the second most common cause of death in the United States, accounting for nearly 1 of every 4 deaths. In 2016, around 1.7 million people were expected to be diagnosed with cancer and 595,690 were expected to die from the disease in the United States [6]. In addition to the devastating effects on patients and their families, the economic costs of cancer are enormous, both in terms of direct medical-care resources for its treatment and in the loss of human capital due to early mortality. For example, the annual cost for cancer care is projected to rise from $104 billion in 2006 to over $174 billion in 2020 [7]. Advances in cancer treatment and improvements in cancer outcomes over the past few decades have been modest, despite significant investment in

cancer research. A great deal of research is invested in improving treatment for advanced disease because cancers detected at advanced stages are far more likely to cause death than those detected while the cancer is still confined to the organ of origin. Most people who develop cancer have advanced disease at the time of diagnosis. For example, 65% of patients with oral cavity and pharynx cancer, 57% of those with colorectal cancer, and 34% of individuals with breast cancer have regional or distant spread of disease at the time of diagnosis. Early detection represents one of the most promising approaches to improve both survival and quality of life of cancer patients because it enables a surgical, curative approach. For the overall population, shifting all cases to early detection would have a significant impact on overall cancer mortality and economic burden [8]. Screening methods that can detect precursor lesions or in situ disease hold even more promise, namely the possibility of eliminating the invasive condition entirely and with it the burden of the disease.

Over 80% of malignancies occur in epithelial surfaces, most of which can be directly visualized [9]. Thus, many current procedures for cancer screening begin with visual inspection of the entire tissue surface at risk under white light illumination, followed by biopsy of highly suspicious tissue regions to make definitive diagnosis of its type and cancerous potential [10]. Due to the heterogeneous morphology and visual appearance of the lesions, biopsy diagnosis may not be representative of the highest pathological grade of a tumor due to the small sampling area [11]. After biopsies, tissue samples are sectioned and stained. Pathologists then examine the specimens under microscopes and make judgments based on observations of cell morphology and colors of different tissue components. This biopsy procedure is time-consuming and invasive. In addition, the interpretation of the histological slides is subjective and can be inconsistent due to intra-observer and inter-observer variation [10] [12]. Noninvasive alternatives have been sought using a number of modalities including Computed tomography (CT), ultrasound, and magnetic resonance imaging (MRI). While the enhanced capabilities of

these imaging systems will continue to play a key role in improving patient care, the financial cost and infrastructure required to establish and maintain them restricts their use largely to regional healthcare centers in industrialized countries. Optical imaging may provide a potential solution to the global need for affordable imaging tools to aid in the early detection and management of cancer [9].

1.1.2 Surgical margin assessment

After cancer detection, surgery is usually limited to tumors detected at an early stage and the likelihood of a favorable outcome decreases significantly once primary surgery is not a treatment option. About two-thirds of patients with squamous cell carcinoma of the head and neck (SCCHN) have advanced stage disease at the time of diagnosis, while the remaining one third of all patients present with stage I or stage II disease [13]. Up to 90% of patients with stage I disease and about 70% of those with stage II disease can be cured by surgery or radiation therapy. Therefore, early detection and subsequent surgical removal is of great importance for the prognosis of cancer patients.

If all the cancer cells are removed by surgery, the patient can be cured of that cancer. The presence or absence of the remaining tumor cells in the area surrounding the resection, also known as the surgical margin, is usually considered one of the strongest predictors of tumor recurrence and survival. Positive margins defined as tumor cells being present at the cut edge of the surgical specimen, are associated with increased local recurrence risk and indicate poor prognoses for patients with head and neck cancer, breast cancer, non-small-cell lung cancer, colorectal cancer, and urogenital tract cancer. In most cases, the poorer outcome as a result of positive surgical margins is not mitigated by salvage surgery (that is, re-excision of the positive margin) or by adjuvant chemotherapy and/or radiotherapy [14]. Thus, improvement in the completeness of tumor removal would benefit patients and might produce significant cost savings [15].

However, visual appearance and palpation is the only way to for a surgeon to differentiate between tumor and normal tissue and consequently determine an adequate tumor-free margin during surgery [16]. This process is subjective and not easily quantifiable. Intraoperative frozen margin evaluation is a commonly used practice to assess tumor margins for head and neck surgery [15]. However, frozen section diagnosis may suffer from errors that occur during sampling and histological interpretation; in addition, the histological processing can take time, which is labor intensive, and extends anesthesia-related risks [17]. Therefore, a way to objectively assess tumor margins during surgery in patients would be of great value.

1.2 Specific Aims

This study proposes to investigate the utility of HSI in combination with machine learning methods for quantitative detection and delineation of head and neck cancer. The ultimate goal is to develop a diagnostic tool for early cancer detection and image-guided surgery of head and neck cancer.

Specific Aim 1: Develop and evaluate image quantification methods for HSI to noninvasively detect cancer in xenograft cancer models. a) To set up a preclinical protocol for HSI in a xenograft cancer model with HSI. b) To develop and validate a pre-processing, feature extraction and dimension reduction, classification, and post-processing method for the differentiation between cancer and healthy tissue and for the estimation of cancer likelihood for each pixel on hyperspectral images. c) To compare the proposed spectral-spatial classification method with conventional spectral classification methods. ***Impact***: This study will demonstrate that (1) the proposed spectral-spatial classification method is superior to conventional spectral classification methods, and (2) the combination of HSI with machine learning methods enables quantitative detection of cancers in a noninvasive manner in animal models.

Specific Aim 2: Develop and validate hyperspectral image classification methods for the detection and delineation of tongue cancer in a chemically-induced tongue carcinogenesis model. a) To design the imaging procedures for the chemical-induced mouse tongue tumor model. b) To develop image classification methods for the distinction of dysplasia, carcinoma in situ (CIS), carcinoma and healthy tissue, and to establish validation methods by reconstruction of a pathological gold standard map of tongue. d) To compare reflectance HSI with autofluorescence and vital-dye fluorescence imaging for tongue cancer detection and delineation. e) Identify the correlation between reflectance spectral signatures with histological features. *Impact*: This study, for the first time, demonstrates the feasibility of HSI to detect cancers from mouse tongues without exogenous contrast agents. The success of this work could provide insights into the pathophysiology underlying the spectral differences of tumor and normal tissue.

Specific Aim 3: Develop and validate a hyperspectral image processing and quantification framework for intraoperative cancer detection in xenograft cancer models. a) To set up experimental methods for cancer surgery and imaging in a subcutaneous cancer model in mice. b) To develop a framework for hyperspectral image processing and quantification, including a set of steps consisting of image pre-processing, glare removal, registration, feature extraction and selection, and classification for intraoperative cancer detection. *Impact*: This work will investigate the feasibility of HSI with fast image classification for head and neck cancer detection during animal tumor surgery.

Specific Aim 4: Develop and validate machine learning-based hyperspectral image quantification methods for the detection of head and neck cancer in fresh human surgical specimen. a) To design study protocols for collecting fresh surgical tissue from patients with primary head and neck squamous cell carcinoma (HNSCC) and for procuring, processing, and imaging the tissue specimen. b) To collect tissue, perform imaging experiments, register histology with hyperspectral images, and to develop and

validate machine learning methods to differentiate cancer from normal tissue using a variety of head and neck cancer sites. c) To compare the performance of reflectance HSI, autofluorescence, and vital-dye fluorescence imaging for cancer diagnosis. *Impact*: This is the first study to use HSI to detect and delineate tumors from a wide variety of head and neck sites including oral cavity, gland (thyroid and parotid), larynx, pharynx, paranasal sinus, and nasal cavity of human patients. The machine learning methods could provide fast and quantitative assessment of tumor margins during surgery.

1.3 Thesis Overview

This thesis investigated the potential of HSI with machine learning methods for the detection and delineation of head and neck cancers in preclinical animal models and fresh surgical specimens from human patients. Figure 1 illustrated the structure of this dissertation, linking the chapters to the four specific aims.

Chapter 2 of this thesis introduced and explaind HSI technology and gives an overview of the literature on HSI hardware, software, and medical applications. This chapter explained HSI technology and gave an overview of HSI hardware, image analysis, and medical applications relevant to this dissertation. Image analysis methods for MHSI were summarized with an emphasis on preprocessing, feature extraction and selection, and classification methods. The section on applications summarized the available literature on cancer diagnosis and surgery guidance. Finally, we concluded with a discussion on the achievements of the past few years and some future challenges.

Chapter 3 of this thesis investigated the use of hyperspectral imaging combined with a spectral-spatial classification method for non-invasive head and neck cancer detection in a subcutaneous cancer model. Hyperspectral reflectance images were acquired from 450 nm to 900 nm with a 2-nm increment from tumor-bearing mice. The hyperspectral imaging and classification method was able to distinguish cancer from normal tissue on hyperspectral images with a sensitivity of 93.7% and a specificity of

91.3% in the animal experiment. The preliminary study demonstrated that HSI has the potential to be applied *in vivo* for noninvasive cancer detection.

Chapter 4 of this thesis evaluated the diagnostic performance of HSI for the detection and delineation of tongue cancer in a chemically-induced tongue carcinogenesis model. Color-coded pathology gold standard maps were reconstructed to match the dorsal surface of the tongue for validation. Multiple machine learning classifiers, including discriminant analysis, ensemble learning, and support vector machines, were evaluated for tongue neoplasia detection with HSI, which were validated by the reconstructed pathological maps. The diagnostic performance of HSI, autofluorescence imaging (AFI), and fluorescence imaging (FI) were compared for neoplasia detection. Prediction maps were generated to display the location and distribution of neoplasia. With the proposed algorithm, *in vivo* HSI achieved an average AUC, sensitivity, and specificity of 0.852, 76.4%, and 81.8%, and *ex vivo* HSI achieved an average AUC, sensitivity, and specificity of 0.862, 77.3% and 80.9%.

Chapter 5 of this thesis proposed a framework for hyperspectral image processing and quantification, which included a set of steps including image preprocessing, glare removal, feature extraction, and ultimately image classification. The framework was tested on images from mice with head and neck cancer. The image analysis extracted multiple spectral features including Fourier coefficients, normalized reflectance, mean, and spectral derivatives for classification. The experimental results demonstrated the feasibility of the hyperspectral image processing and quantification framework for cancer detection during the animal tumor surgery in a challenging setting where sensitivity can be low due to a modest number of features present but potential for fast image classification can be high. This HSI approach may have the potential application in tumor margin assessment during image-guided surgery, where speed of assessment may be a dominant factor.

Chapter 6 of this thesis presented a proof-of-concept study to determine the feasibility of using HSI for identifying and differentiating tumor from normal tissue from a variety of anatomic sites of head and neck cancer patients. This chapter showed that tumors could be differentiated from normal tissues using HSI with both intra- and inter-patient classification methods. The cancerous regions delineated by the automated classification methods were validated by pathological diagnosis. The experimental results suggested that HSI combined with machine learning techniques could be developed for the rapid, accurate, and objective assessment of the surgical margins during head and neck cancer surgery.

Chapter 7 of this thesis summarized the contribution of this paper and future directions of the research.

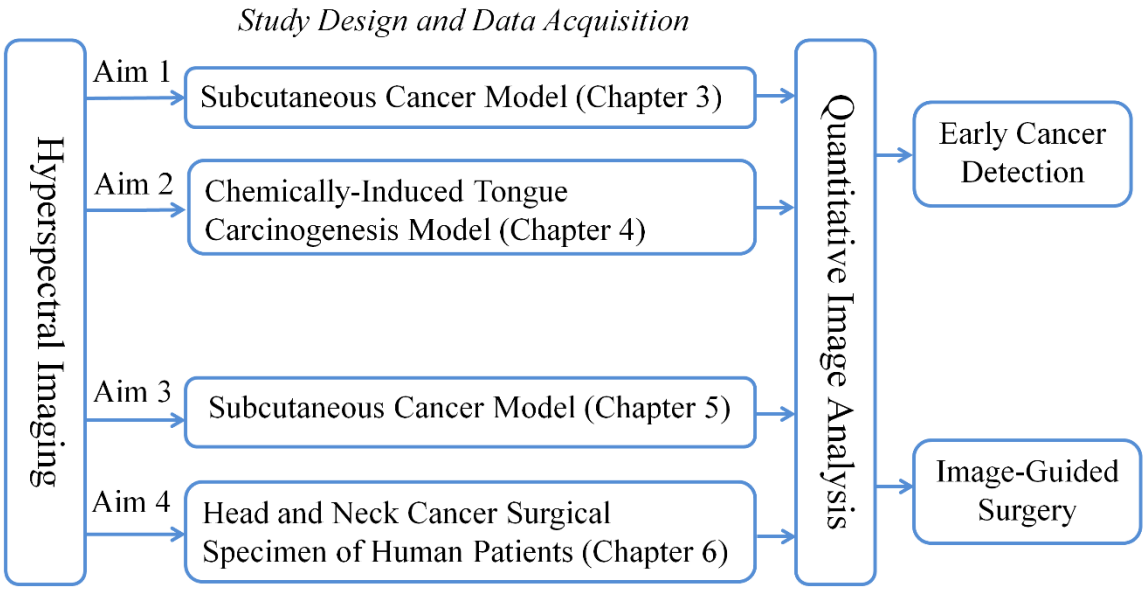


Figure 1: Structure of dissertation.

CHAPTER 2

MEDICAL HYPERSPECTRAL IMAGING: A REVIEW

Hyperspectral imaging originated from remote sensing and has been explored for various applications by NASA [18]. With the advantage of acquiring two-dimensional images across a wide range of the electromagnetic spectrum, HSI has been applied to numerous areas including archaeology and art conservation [19] [20], vegetation and water resource control [21] [22], food quality and safety control [23] [24], forensic medicine [25] [26], crime scene detection [27] [28], and biomedicine [29] [30].

As an emerging imaging modality for medical applications, HSI offers great potential for non-invasive disease diagnosis and surgical guidance. Light delivered to biological tissue undergoes multiple scattering events from the inhomogeneity of biological structures and absorption primarily in hemoglobin, melanin, and water as it propagates through the tissue [31] [32]. It is assumed that the absorption, fluorescence and scattering characteristics of tissue change during the progression of disease [33]. Therefore, the reflected, fluorescent, and transmitted light from tissue captured by HSI carries quantitative diagnostic information about tissue pathophysiology [33] [34] [35] [33, 36]. In recent years, advances in hyperspectral cameras, image analysis methods, and computational power make it possible for many exciting applications in the medical field.

In the following, we aim to introduce and explain MHSI technology and to give an overview of the literature on MHSI hardware, image analysis, and medical applications relevant to this dissertation. For a broader technology overview and more examples of medical application, readers are referred to our review paper [37].

2.1 Medical Hyperspectral Imaging Systems

HSI is a hybrid modality that combines imaging and spectroscopy. By collecting spectral information at each pixel of a two-dimensional (2D) detector array, HSI generates a three-dimensional (3D) dataset (x, y, λ), comprised of two spatial dimensions (x, y) and one spectral dimension (λ), known as a hypercube. As shown in Figure 2, each plane of the hypercube represents a grayscale image at a particular wavelength, and intensities over all the spectral bands form a spectral signature for each pixel of the hypercube. The spectral signature of each pixel characterizes the composition of that pixel and correlates with the biochemical and morphological changes in tissue. So hyperspectral images, which contain spectral signature at each image point, can be analyzed to identify various pathological conditions. With the spatial information, the source of each spectrum can be located, which makes it possible to correlate the light interactions with pathology in a much larger area than conventional spectroscopy.

HSI generally covers a contiguous portion of the light spectrum with more spectral bands (up to a few hundred) and higher spectral resolution than multispectral imaging (such as RGB color cameras). Therefore, HSI has the potential to capture the subtle spectral differences under different pathological conditions, while multispectral imaging may miss significant spectral information for diagnostics. The difference between a hypercube and an RGB color image is illustrated in Figure 2. Among all the MHSI systems investigated in the literature, the majority of the systems are prototypes consisting of off-the-shelf components rather than commercialized systems.

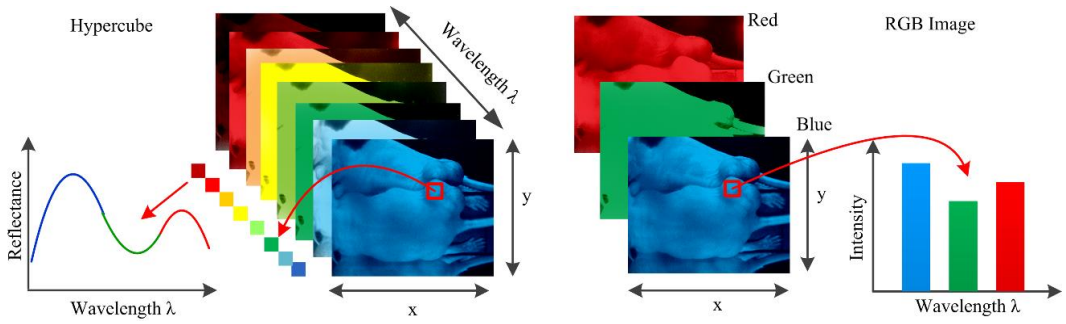


Figure 2: Comparison between a hypercube and RGB image. A hypercube is a three dimensional dataset comprised of 2D images of each wavelength. The lower left is the reflectance curve (spectral signature) of a pixel in the image. RGB color images only have three image bands: red, green, and blue respectively. The lower right is the intensity curve of a pixel in the RGB image.

HSI systems consist of a light source, wavelength dispersion devices, area detectors, and a computer platform for storage, display, and processing of imaging data. Dispersive devices are the core element of an HSI system, which are either located between the light source and the sample for excitation wavelength selection, or between the sample and the detector arrays for emission wavelength dispersion. There are many types of optical and electro-optical dispersive devices which can perform spectral dispersion or selection in HSI systems. The commonly used dispersive devices in the literature can be divided into three classes; (1) Monochromators: prism and diffraction grating, (2) Optical bandpass filters: fixed filter or tunable filters, (3) Single shot imagers such as a computer-generated hologram (CGH). A detector array or detector FPA is an assemblage of individual detectors located at the focal plane of an imaging system [38]. The most widely used detector arrays in the literature are Charge-Coupled Devices (CCDs) because of their high quantum yield and very low dark current. CCDs consist of many photodiodes which are composed of light-sensitive materials such as silicon (Si), indium gallium arsenide (InGaAs), indium antimonite (InSb), and mercury cadmium telluride (HgCdTe). Based on the spectral response of these materials, the working

wavelength range of CCDs varies from UV to NIR. Cooling CCDs can lower the operating temperature of the detectors and therefore reduce dark-current noise. Silicon CCDs are mostly used in the VIS and NIR regions in MHSI systems [30] [39] [40] [41] [42] [43], due to their high resolution, and acceptable quantum efficiency in the spectral range.

Depending on how spectral and spatial information is acquired, HSI generally involves two scanning methods: spatial scanning and spectral scanning. Spatial scanning methods generate hyperspectral images by acquiring a complete spectrum for each pixel in the case of whiskbroom (point-scanning) instruments, or line of pixels in pushbroom (line-scanning) instruments, and then spatially scanning through the scene. Whiskbroom and pushbroom HSI do not provide live displays of spectral images, which are calculated from the spectra after the completion of the spatial scanning of the corresponding area. Spectral scanning methods, also called staring or area-scanning imaging, involve capturing the whole scene with two-dimensional detector arrays in a single exposure, and then stepping through wavelengths with an imaging monochromator to complete the data cube. The use of staring HSI to build the hypercube has the advantage of displaying live spectral images, which is essential for aiming and focusing [33]. Area-scanning imaging is suitable for stationary applications such as samples under a hyperspectral microscope.

Hyperspectral Imaging System Used in This Dissertation

In this dissertation, all the hyperspectral dataset were obtained with a wide-field staring imaging called CRI Maestro *in-vivo* imaging system This system mainly consists of a flexible fiber-optic lighting system, a solid-state liquid crystal tunable filter (LCTF, bandwidth 20 nm) as the wavelength dispersion device, a spectrally optimized lens, and a 12-bit high-resolution charge-coupled device (CCD) as the area detector. A

Cermax-type, 300-Watt, Xenon light source is used as white light excitation, which spans the electromagnetic spectrum from 450 nm to 800 nm. The interior near-infrared light source can be used to extend the spectral range of excitation up to 950 nm. Four fiber-optic, adjustable illuminator arms yield an even light distribution to the subject. The reflected light from the surface of the subject is split into a series of narrow spectral bands by the LCTF. The LCTF can be electronically controlled through application of voltage, thus allowing for the selection of one band at a time. At each spectral scanning step, a full spectral image is recorded by a 12-bit charge-coupled device (CCD), interfaced with a computer platform for storing and displaying the live spectral images of the subject. The active light sensitive area of the CCD is 1,392 pixels in the horizontal direction and 1,040 pixels in the vertical direction. The hyperspectral camera can simultaneously acquire full datasets from as many as three mice in only a few seconds. For image acquisition, the wavelength setting can be defined within the range of 450-950 nm (with the interior near-infrared light on) with 2 nm increments (or 5 nm, 10 nm, etc.); therefore the data cube collected was a 3D array of the size 1,040 × 1,392 × N. The number of the spectral images N is determined by the wavelength range and the increments as chosen by the user. The field of view (FOV) is from 3.4 × 2.5 cm to 10.2 × 7.6 cm. One advantage of the staring hyperspectral imaging technique like Maestro is that measurements can be performed remotely without contact with the subject and without any moving parts in the system. Non-contact detection can be particularly useful in clinical settings where fiber-optic probes may contribute to a number of problems, including contaminating sterile fields, altering regional tissue perfusion (through probe-induced pressure), and increasing measurement uncertainty that is due to irreproducible fiber-tissue coupling. The drawback of such a design is that it is not possible to eliminate crosstalk between adjacent tissue pixels due to tissue scattering.

2.2 Rationale for Cancer Detection with Hyperspectral Imaging

The propagation of light within tissue is a significant issue affecting medical applications and in the development of diagnostic methods. Therefore, this section is dedicated to a brief review of the light tissue interaction mechanisms, optical processes involved in HSI, and useful diagnostic information provided by HSI.

Light entering biological tissue undergoes multiple scattering and absorption events as it propagates across the tissue [44]. Biological tissues are heterogeneous in composition with spatial variations in optical properties [45]. Scattering occurs where there is a spatial variation in the refractive index [45]. In cellular media, the important scatterers are the subcellular organelles, with their size running from <100 nm to 6 μm. For example, mitochondria are the dominant scatterers among the organelles. The structure of a lipid membrane and lipid folds running inside gives mitochondria a high optical contrast to the surrounding cytoplasm and produces the observed strong scattering effects. The shape and size of the cells vary among different tissue types with dimensions of a few microns and larger [45]. The scattering properties of support tissues composed of cells and extracellular proteins (elastin and collagen, etc.) are caused by the small-scale inhomogeneities and the large-scale variations in the structures they form.

The penetration depth of light into biological tissues depends on how strongly the tissue absorbs light. Most tissue are sufficiently weak absorbers, permitting significant light penetration within the therapeutic window, ranging from 600 to 1300 nm [45]. Within the therapeutic window, scattering is dominant over absorption, so the propagating light becomes diffuse. Tissue absorption is a function of molecular composition. Molecules absorb photons when the photon energy matches an interval between internal energy states, and the transition between quantum states obeys the selection rules for the species. During absorption processing, transitions between two energy levels of a molecule that are well defined at specific wavelengths could serve as a spectral fingerprint of the molecule for diagnostic purposes [45] [46]. For example, absorption spectra characterize the concentration and oxygen saturation of hemoglobin,

which reveals two hallmarks of cancer: angiogenesis and hypermetabolism [32]. Tissue components absorbing light are called chromophores. Some of the most important chromophores for visible wavelengths are blood and melanin. The primary absorbers for UV wavelengths are protein and amino acids, while the important absorbing chromophore for infrared (IR) wavelengths is water [47].

Light absorbed by tissue constituents is either converted to heat or radiated in the form of luminescence, including fluorescence and phosphorescence [45] [48] [34]. Fluorescence that originates from endogenous fluorescent chromophores is also called autofluorescence. Incident light, typically in the UV or VIS region, excites the tissue molecules and induces fluorescence emission. The majority of the endogenous fluorophores are associated with the structural matrix of tissue or with various cellular metabolic pathways [45] [49]. The most common fluorophores in the structural matrix are collagen and elastin, while the predominant fluorephores involved in cellular metabolism are nicotinamide adenine dinucleotide (NADH), flavin adenine dinucleotide (FAD), and lipopigments [50]. These intrinsic fluorophores exhibit different strengths and cover spectral ranges in the UV and VIS regions. For example, fluorescence from collagen or elastin using excitation between 300 and 400 nm shows broad emission bands between 400 and 600 nm, which can be used to distinguish various types of tissues, e.g., epithelial and connective tissue [51]. Cells in different disease states often have different structures, or undergo varying rates of metabolism, which result in different fluorescence emission spectra. Therefore, fluorescence imaging makes it possible to investigate tissues for the diagnosis of diseases in real time without administrating exogenous fluorescent agents [45]. Various exogenous fluorophores have also been created and studied for biological diagnostics using HSI [50], but this review will mainly discuss the intrinsic fluorescence.

Incident light can be directly reflected on the surface of the tissue, or be scattered due to random spatial variations in tissue density (membranes, nuclei, etc.) and then be

reemitted to the tissue surface [48]. Light becomes randomized in direction due to multiple scattering events, known as diffuse reflectance, which provide information about scattering and absorbing components deep within the tissue [52]. The measured reflectance signal represents light that has sampled a variety of sampling depths within the tissue, and is therefore a weighted average measure of the properties over a certain volume of tissue [52]. Knowledge of the origin of the scattering and absorption signal would facilitate accurate modeling and interpretation of the reflectance data. The reflectance signal measured from epithelial tissue is determined by the structural and biochemical properties of the tissue; therefore, changes in optical properties can be used to noninvasively probe the tissue microenvironment [52]. Alterations in tissue morphology including hyperplasia, nuclear crowding, degradation of collagen in the extracellular matrix by matrix metalloproteinases (MMPs), and increased nuclear/cytoplasmic ratio, which are associated with disease progression, can affect the scattering signals. As diseases progress, hemoglobin absorption may be affected by angiogenesis and tissue hypoxia. Therefore, changes in the disease states should lead to corresponding changes in the patterns of the light reflected from the tissue, although with potential for high variability.

Head and neck cancers (HNC) mainly include tumors in the lips, oral cavity, nasal cavity, oropharynx, hypopharynx, and larynx. Most of HNCs are squamous cell carcinomas (SCC) that originate from the epithelial region. The vast majority of epithelial cancers are preceded by precancerous changes that affect both the surface epithelium and deeper stroma. Pre-cancers are accompanied by local metabolic and architectural changes at cellular and subcellular level, such as changes in the nuclear-to-cytoplasmic ratio of cells, mean nuclear diameter, nuclear size distribution, nuclear refractive index, and chromatin texture. These changes affect the elastic scattering properties of tissue. Pre-cancers are characterized by increased metabolic activity, which affects mitochondrial fluorophores NADH and FAD, and changes the fluorescence properties of tissue [53].

For instance, the clinical appearance of the oral mucosa varies depending on the distribution of superficial blood vessels, the type and thickness of epithelium, components of the submucosa, and presence of pigmentation and appendages, etc. These biochemical and morphological features can change during disease progression, so the altered spectral fingerprint captured by hyperspectral imaging carries diagnostic information which can be used to differentiate cancer from healthy tissue.

2.3 Hyperspectral Image Analysis

A hyperspectral dataset contains rich spectral-spatial information for disease diagnostics. For example, a HSI system operating in the wavelength range of 450-900 nm with a 2-nm interval, with an image size of 1040 × 1392 pixels in the spatial dimension, will generate 1.45 million spectra in one hypercube, each with 226 data points. The ability to pull out important diagnostic information from the large volume of a dataset is the key to fully exploit the diagnostic potential of hyperspectral imaging. Hypercubes with high spatial and spectral resolutions may potentially contain more diagnostic information. The major challenges for automatic analysis of hyperspectral data with high spatial and spectral dimensions involve several aspects: 1) high data redundancy due to high correlation in the adjacent bands, 2) intra- and inter- subject variability of hyperspectral signatures, and 3) curse of dimensionality [54].

As a combination of spectroscopy with imaging, HSI measures the light intensity as a function of both wavelength and location. In the spectroscopy domain, a fully resolved spectrum at each pixel location can be recorded. In the image domain, the data set includes a full image at each wavelength band. Past spectroscopic analysis methods for fiber-optic-based measurements of diffuse reflectance spectra mainly include analytical approximation of the light transport equation, Monte Carlo model, and empirical method, which allow the extraction of the absorption and scattering coefficients of tissue by minimizing the difference between the measured and the fitted reflectance

spectrum. These methods have the advantage of directly linking the optical parameters to the disease state of tissue and providing insight into the underlying mechanisms, thus enabling better interpretation of hyperspectral imaging data for cancer detection.

Image analysis with machine learning techniques provides an indirect way to link the spectral-spatial features with the disease states. These methods exploit the full spectral information, and make no assumption about tissue structure and composition, which can be easily adapted to different imaging modalities and tissue types. Although the analysis methods for hyperspectral images of the earth surface in the field of remote sensing have been intensively investigated, their development and application in the medical domain lag far behind. The basic steps for hyperspectral image classification generally involve pre-processing, feature extraction and feature selection, and classification. A probability map depicting pathological status of the tissue can be generated by the classification methods, which would enable the detection and grading of pathology, provide image-guided surgical treatment, and facilitate objective follow-up. In the following, each step of the hyperspectral image analysis method will be discussed.

2.3.1 Data Preprocessing

HSI preprocessing mainly involves reflectance normalization and noise removal: Reflectance normalization converts or normalizes hyperspectral radiance observations to percent reflectance [55] [56] values that describe the intrinsic properties of biological samples. Such normalization also reduces system noise and image artifacts arising from uneven surface illumination or large redundant information in the subbands of hyperspectral imagery, and better prepares data for further analysis. CCD arrays used in HSI systems generally have dark current even without light shining on it, where dark current is dependent on temperature and is proportional to integration time. To convert raw intensity into reflectance, white reference and dark images are taken before acquiring sample images. The white reference image is taken with a standard reflectance surface

placed in the scene, and the dark current measured by keeping the camera shutter closed. Currently, the widely-used standard reflectance surface is the NIST (National Institute of Standards and Technology) certified 99% Spectralon White Diffuse Reflectance target. The raw data was then corrected using the following equation [57]:

$$I_{ref}(x, y, \lambda) = \frac{I_{raw}(x,y,\lambda) - I_{dark}(x,y,\lambda)}{I_{white}(x,y,\lambda) - I_{dark}(x,y,\lambda)} \quad (1)$$

Where $I_{ref}(x, y, \lambda)$ is the normalized reflectance value at pixel location (x, y) and wavelength band λ. $I_{raw}(x, y, \lambda)$ is the corresponding raw radiance value, and I_{dark} and I_{white} are the dark current and the white reference intensity of the given pixel, respectively. To smooth the spectral signature and further reduce noise effects, a Gaussian filter [58] and a Savitsky-Golay linear filter [59] were used in the literature. In addition, *in vivo* imaging may suffer from motion artifacts from the subjects, thus requiring image registration to account for tissue movement. Image registration finds a geometric transformation of multiple images of the same scene taken at different wavelengths. The correspondence between the images is maximized when an image pair is correctly aligned. To obtain accurate spectral information for each pixel, image registration may be necessary to spatially align all spectral band images within one hypercube or between different hypercubes. Kong *et al.* [58] utilized mutual information (MI) as a metric for searching the offset of the band images along the horizontal axis, and an image pair with maximum MI shows the best match between a reference image and an input image. Each band image was spatially co-registered to eliminate the spectral offset caused during the image acquisition procedure. Panasyuk *et al*. [60] performed image registration as a preprocessing step to account for slight motions during the imaging of anesthetized mice.

2.3.2 Feature Extraction and Selection

In hyperspectral datasets, each pixel (x, y) can be represented by an N-dimensional vector with the normalized reflectance R at each individual wavelength λ_i:

$[R(x, y, \lambda_1), R(x, y, \lambda_2), \cdots, R(x, y, \lambda_N)]$ (N is the number of the spectral bands), which is the most common way of feature construction. Such pixel-based representation has been widely used for hyperspectral image processing tasks. This method treats hyperspectral data as unordered listings of spectral measurements without particular spatial arrangement [61], which may result in a salt-and-pepper look for the classification map. Therefore, feature extraction methods incorporating both spatial and spectral information have been investigated intensively in the remote sensing area to improve classification accuracy. Recent advances of spatial-spectral classification have been summarized in [62].

For hyperspectral datasets, a larger number of spectral bands may potentially make the discrimination between more detailed classes possible. But due to the curse of dimensionality, the use of too many spectral bands during classification may decrease the classification accuracy [54]. Therefore, it is important to perform feature extraction and selection to obtain the most relevant information from the original data and represent that information in a lower dimensional space. The most widely used dimensionality reduction method for medical hyperspectral dataset analysis is principle component analysis (PCA). PCA reduces redundant information in the bands of hyperspectral imagery while preserving as much of the variance in the high dimensional space as possible. PCA is optimal in the sense of minimizing the mean square error. However, PCA transforms the original data to a subspace spanned by eigenvectors, which makes it difficult to interpret the biological meaning after transformation. Several PCA variants, such as minimum noise fraction (MNF) [43] and independent component analysis (ICA) [63] are also used for feature extraction and dimensionality reduction.

While traditional dimension reduction methods such as PCA and LDA project the original features into a lower dimensional space, feature selection techniques select a subset of the original features without a transformation. Thus, they preserve the original semantics of the variables, offering the advantage of interpretability.

2.3.3 Classification

The goal of supervised classification is to learn the underlying patterns of training data and build predictive models to accurately differentiate cancer from normal tissue in hyperspectral images. Commonly used classifiers for hyperspectral image classification include linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), ensemble LDA, RUSBoost, random forests, RUSBoost, linear support vector machine (SVM), kernel SVM with radial basis function (RBF), and artificial neural networks (ANN). In the following, we briefly describe the characteristics of these classifiers.

Discriminant Analysis (LDA and QDA)

Discriminant analysis aims to find the projections of the original high dimensional space in a lower dimensional space where the class separation is maximized [64]. Both LDA and QDA classification models assume that the data has a multivariate normal distribution [64] [65]. The difference between them is that the LDA model assumes the same covariance matrix for each class with varying means, while QDA assumes an individual mean and covariance matrix for each class [65] [66]. LDA is advantageous in several aspects. Firstly, it does not require tuning of free parameters. Secondly, it has been shown to perform well in classification tasks, even though the assumption of a common covariance matrix among groups and normality are violated [67]. Lastly, the decision hyperplane for binary classification obtained by SVM is found to be equivalent to the solutions by LDA on the set of support vectors [68]. LDA has been reported to be the best among other classifiers (K-nearest neighbor (KNN), decision tree (DT), QDA, ensemble KNN, ensemble-DT, etc.) in classifying multispectral images of burn wound tissues [69].

Ensemble LDA

Ensemble learning methods build a high-quality ensemble predictor by combining results from many weak learners such as decision trees and discriminant learners. Ensemble LDA improves the accuracy of LDA with random subspace ensembles [70], which involve training multiple LDA learners with randomly selected subsets of features without replacement and classifying the testing data by taking the average of the scores predicted by the weak learners. Two parameters in this model need to be tuned for optimal classification performance: the number of feature dimensions to sample in each learner and the number of learners in the ensemble.

RUSBoost

RUSBoost is an ensemble learning method that is especially effective at classifying imbalanced dataset [71]. Random Under Sampling (RUS) is a commonly used data sampling technique that randomly removes examples from the majority class in order to adjust the class distribution of the training data set. AdaBoost is a popular boosting technique that has been shown to improve the classification performance of weak learners. RUSBoost combines RUS with AdaBoost techniques to improve classification performance for skewed data. This method first generates a balanced training dataset for each weak learner in the ensemble by taking N, the number of instances in the minority class, as the basic unit for sampling and selecting a subset of the majority classes with N instances, and then follows the boosting procedure to iteratively create an ensemble of weak learners and make predictions by the weighted vote of individual weak learners. Here, a decision tree is chosen as the weak learner. We control the depth of a decision tree by adjusting the number of observations per leaf node and the maximal number of branch node splits per tree. In addition, we also select the optimal number of tree learners in the ensemble by cross validation.

Random Forest

A random forest is a classification algorithm that uses an ensemble of decision trees [72]. Random forest combines bagging and random feature selection to yield an ensemble that can achieve both low bias and low variance. Each tree is grown on a training set randomly drawn with replacement from the original dataset. In each bootstrap training set, about one-third of the instances are left out for out-of-bag estimates. At each node of the tree, a random subset of features is selected from the input features to split on. The final classification decision is taken by averaging the class probabilities calculated by all produced trees. Generally, two parameters need to be set in order to produce the forest trees: the number of trees to be generated and the number of features to be selected. Random forest classifier has been widely used in the remote sensing field due to its classification accuracy [73]. Our previous work also demonstrated the successful use of random forests for tongue cancer detection [74].

Support Vector Machines (SVMs)

SVM is a kernel-based machine learning technique which has been widely used in the classification of hyperspectral images [75] [58] [76] [77] [78] [79] [80] [81] [82] [83] [84] [75] [85]. Due to its strong theoretical foundation, good generalization capability, low sensitivity to the curse of dimensionality [86] and ability to find global classification solutions, SVM is usually preferred by many researchers over other classification paradigms. Given training vectors $x^i \in \mathbb{R}^N, i = 1, 2, \cdots, M$ in two classes, and an indicator vector $\mathrm{y} = [\mathrm{y}^1, \mathrm{y}^2, \cdots, \mathrm{y}^M]^T \in \mathbb{R}^M$ such that $y^i \in \{1, -1\}$, C-Support Vector Classification [87] [88] solves the following primal optimization problem:

SVM has been proved to perform well for classifying hyperspectral data [75]. In the processing of medical hyperspectral data, SVM has also been explored for various classification tasks. Melgani *et al.* [75] investigated the effectiveness of SVMs in the classification of hyperspectral remote sensing data. It was found that SVMs were much more effective than radial-basis function (RBF) neural networks and the K-Nearest

Neighbor (KNN) classifier in terms of classification accuracy, computational time, and stability to parameter settings. Kong *et al.* [58] chose a Gaussian RBF kernel as the kernel function for SVM, and learned the SVM parameters from 100 training samples chosen randomly from each of the normal and tumor classes. For testing, 2036 normal and 517 tumor samples were used. Experimental results showed that the spatial filtering enhanced the performance, which resulted in an overall accuracy of 86% while the use of the original data had an accuracy of 83%.

In our group, we used SVMs for various tissue classification tasks. In [78], Akbari *et al.* extracted and evaluated the spectral signatures of both cancerous and normal tissue and used least squares SVMs to classify prostate cancer tissue in tumor-bearing mice and on pathology slides. In [79], they created a library of spectral signatures for different tissues and discriminated between cancerous and non-cancerous tissues in lymph nodes and lung tissues with SVMs. In [80], Akbari *et al*. constructed a library of spectral signatures from hyperspectral images of abdominal organs, arteries, and veins, and then differentiated between them using SVMs. In [81], they utilized least squares kernel SVMs to classify normal tissues and tumors based on their standard deviation and normalized difference index of spectra signature. In this thesis, we choose the LIBSVM software package [88] for both linear and kernel SVMs.

Artificial Neural Networks

Neural Network is another supervised classification method that has been adopted by many researchers [89] [90] [91] [92], due to its non-parametric nature and arbitrary decision boundary. Multilayer perceptron (MLP) is the most popular type of neural network in image classification [89]. It is a feed-forward network trained by the back-propagation algorithm. Monteiro *et al.* [55] implemented both single-layer perceptron (SLP) and MLP as supervised classifiers. The MLP notably generated the clearest

visualization of the calendar's number under the blood. Although the SLP was also able to learn a good visualization, the output presented more noise.

2.4 Medical Applications

HSI has tremendous potential in disease detection and image-guided surgery because it is able to detect biochemical changes due to disease progression such as cancer cell metabolism [35] [93] [94]. HSI is able to deliver nearly real-time images of biomarker information, such as oxyhemoglobin and deoxyhemoglobin, and assess tissue pathophysiology based on the spectral characteristics of different tissue [60]. In the literature, a variety of studies have used HSI techniques to augment existing diagnostic methods or to provide more efficient alternatives. HSI has been applied to a wide range of medical problems, such as cardiac disease, retinal disease, diabetic foot, shock, cardiac disease, ischemic tissue, skin burn, retinal disease, diabetes, kidney disease, the diagnosis of hemorrhagic shock [95] [96], the assessment of peripheral artery disease [97], early detection of dental caries [98], fast characterization of kidney stone types [90], and the detection of laryngeal disorders [99]. Cancer detection and Image-guided surgery are two major applications for MHSI.

2.4.1 Cancer Detection

The rationale for cancer detection by optical imaging lies in the fact that biochemical and morphological changes associated with lesions alter the absorption, scattering, and fluorescence properties, and therefore the optical characteristics of tissue can in turn provide valuable diagnostic information. For example, optical absorption can reveal angiogenesis and increased metabolic activity by quantifying the concentration of hemoglobin and oxygen saturation [100]. Kortum *et al*. [101] used optical spectroscopy to detect neoplasia and reported that (i) the increased metabolic activity affects mitochondrial fluorophores and changes the fluorescence properties in pre-cancerous

tissue, and (ii) fluorescence and reflectance spectra contain complementary information that was useful for pre-cancer detection.

Compared to optical spectroscopy which measures tissue spectra point by point, HSI is able to capture images of a large area of tissue, and has exhibited great potential in the diagnosis of cancer in the cervix [102] [103] [35] [93], breast [104] [60], colon [105] [106] [107] [108] [109] [110] [111] [112] [83], gastrointestine [113] [114], skin [115] [116], ovarian [117], urothelial carcinoma [118], prostate [78], esophageal [119], trachea [120], oral tissue [121] [122] [123] [36], tongue [56], lymph nodes, [124] and brain [125]. HSI cancer studies have been performed in the following major aspects: (i) recognizing protein biomarkers and genomic alterations on individual tumor cells *in vitro* [126]; (ii) analyzing the morphological and structural properties of cancer histological specimens to classify the cancer grades; (iii) examining the tissue surface to identify pre-cancerous and malignant lesions *in vivo*; (iv) measuring the tissue blood volume and blood oxygenation to quantify the tumor angiogenesis and tumor metabolism.

Head and Neck Cancer

Most head and neck cancer (HNC) originate from the epithelial region in areas such as the lips, oral cavity, nasal cavity, oropharynx, hypopharynx, and larynx. Therefore, it is possible to detect the cancerous tissue despite the limited penetration depth of HSI,. Oral cancer is a subtype of HNC located in the oral cavity, which is commonly examined by visual inspection and palpation of the mouth. However, this visual screening method depends heavily on the experience and skills of the physicians. Roblyer *et al.* [121] [122] reported the use of a multispectral digital microscope (DMD) for the detection of oral neoplasia in a pilot clinical trial. The proposed DMD was a multimodal imaging method which combines the fluorescence, narrow-band (NB) reflectance, and orthogonal polarized reflectance (OPR) modes. They observed decreased blue/green

autofluorescence and increased red autofluorescence in the lesions and increased visibility of vasculature with NB and OPR imaging.

2.4.2 Surgical Guidance

The success of surgery highly depends on a surgeon's ability to see, feel, and make judgments to identify the lesion and its margins [127]. MHSI holds the potential to extend a surgeon's vision at the tissue, cellular, and molecular levels. The ability of MHSI as an intra-operative visual aid tool has been explored in many surgeries.

First, MHSI could facilitate residual tumor detection during cancer surgery [60]. Surgery remains the foundation of cancer treatment, with the central objective of maximizing the removal of the tumor, without harming adjacent normal tissue. However, cancerous tissue is often indistinguishable from healthy tissue in the operating room, and residual tumors that were not apparent to the surgeon at the time of the procedure were often found at the margin of the resected specimen, which leads to the high mortality rates from recurrent tumors. Therefore, intraoperative assessment of surgical margin is critical for complete resection. For example, Panasyuk *et al.* [60] successfully detected residual tumors of 0.5-1.0 mm intentionally left in the operative bed (see Figure 7) during an intraoperative experiment using MHSI in a rat breast-tumor model. A sensitivity of 89% and a specificity of 94% for the detection of residual tumors, comparable to that of histopathological examination of the tumor bed, were reported. With the aid of MHSI, more extensive resection and more effective biopsy locations may be identified. The complete resection of tumor tissue and the conservation of normal tissue may improve surgery outcome, the preservation of organ function, patient satisfaction, and quality of life.

Second, MHSI could monitor the tissue oxygen saturation during surgery. Tissue blood flow or oxygenation is a positive indicator of viable tissue, which might be otherwise sacrificed when removing tumor with little guidance. It has been shown in [128]

that HSI could monitor the tissue at a rate of 3 frames per second, and thus could detect dynamic changes in blood flow and capture unexpected events during surgery.

Finally, MHSI could enable the visualization of the anatomy of vasculatures and organs during surgery. MHSI has the capacity of real-time imaging, which enables the surgeon to make or confirm diagnosis and evaluate surgical therapy in an ongoing fashion in the operation room [127].

Overall, MHSI has been explored in surgery such as mastectomy [60], gall bladder surgery [129], cholecystectomy [130] [131], nephrectomy [132] [133], renal surgery [134] [135] [128] [132] [136], abdominal surgery [80], and intestinal surgery [137].

2.5 Discussion

Over the past two decades, various studies have shown the exciting potential of HSI techniques in medical applications. MHSI is a noninvasive, and non-ionizing technology, which provides a quantitative way of solving medical problems and it may change the medical world in many ways. With the application of MHSI in the exploration of anatomy, physiology, and pathology, human vision has been extended into IR and near-IR wavelength regions. Due to the noninvasive nature, MHSI can be used for optical biopsy which involves *in vivo* diagnosis of tissue without the need for sample excision and processing [138]. Blood volume is generally considered to increase during angiogenesis, and changes in blood oxygenation can be correlated with tumor metabolic activities [42]. Therefore, MHSI can be employed to map the spatial and temporal relationship of the data, and fully grasp the significance of blood oxygen delivery and hypoxia at microvascular levels during tumor growth and angiogenesis [76]. MHSI is also able to visualize chemical contents of vessels and organs, and monitor tissue blood volume and oxygenation during surgery. The use of MHSI does not require the introduction of agents, which is advantageous compared to imaging techniques that

require contrast agents. Moreover, MHSI is able to provide us with real-time data interactively [127], which enables its usage during surgical procedures.

However, the application of MHSI is limited because it examines only areas of tissue near the surface. The optical penetration depth is defined as the tissue thickness that reduces the light intensity to 37% of the intensity at the surface. For a typical person, the optical penetration depth is 3.57 mm at 850 nm and 0.48 mm at 550 nm. While spectral signatures have little dependence on skin temperature over the NIR region, measured radiance in the thermal infrared (8μm-12μm) has a strong dependence on skin temperature [139]. MHSI can also be limited by the cost of HSI imaging systems and by the ability to extract relevant information from large data sets.

HSI combines spectroscopy with imaging, capturing both the spectral and spatial information of biological samples under investigation and providing spatial mapping of parameters of interest in a noninvasive manner. Spectroscopy is a point-measurement method which only measures a limited number of points, so that the derived optical properties may be biased by local tissue inhomogeneities, and important diagnostic information could be missed. Pressure caused by the contact probe may also affect the optical properties due to the altered local blood content, etc. [140]. Although spectroscopy has been explored extensively for probing molecular, cellular and tissue properties [141] [142] [143] and characterizing correlation of tissue parameters with disease state [144], such fundamental research has not been investigated vigorously in HSI. Therefore, fundamental research about the biological rationale of MHSI is necessary, and spectroscopy can be used to validate HSI systems. It was argued that cross talk between spatial locations could occur when extending to HSI, and the information extracted from one location might be influenced by neighboring locations [145]. Martin *et al*. [146] compared the average hyperspectral fluorescence over an area with a value

obtained for one point on the tissue surface obtained by spectroscopy. They found that the major peaks were consistent between the HSI data and spectroscopic data.

HSI can measure significant amounts of spectral information from a large area of tissue. Most literature reported the feasibility of a certain MHSI application without in-depth analysis of the image data obtained. Some results may suffer from a lack of generality because the image datasets are usually constrained to a specific instrument. Therefore, accessible, accurate and up-to-date spectral databases of tissues, cells and molecules for various diseases are needed in order to offer a valuable tool for disease diagnosis and treatment. For example, each subtype of renal tumors such as clear cell, chromophobe, oncocytoma, papillary, and angiomyolipoma can have different morphological and molecular characteristics and thus lead to the differences in spectra signature. Therefore, a spectral library for renal tumors may be able to provide the reference spectra in order to aid the interpretation of hyperspectral images. Furthermore, advanced machine learning methods are to be investigated in order to fully utilize the abundant spectral and spatial information provided by MHSI.

2.6 Challenges in HSI

During the past two decades, HSI technology has undergone fast development in terms of hardware and systems, and has found numerous applications in medical domain. HSI can be easily adapted to several conventional techniques, such as microscopy, fundus camera, colposcopy, laparoscope, etc. to augment the diagnostic performance of existing medical imaging techniques. Three major challenges confront the development and applications of HSI technology.

The first challenge is the acquisition of high resolution HSI datasets in video rates. Real-time acquisition will facilitate intraoperative imaging of the organs, tissues, cells, and molecular biomarkers of interest. Higher spectral and spatial resolution and a larger database of tissue spectra will provide more spatial and spectral information and

may potentially capture more subtle spectral and spatial variations of different tissue types.

The second challenge involves the fast processing of the vast amount of dataset acquired by HSI, including the extraction of the high-quality diagnostic information, and generation of a quantitative map of different tissue types as well as disease-specific endogenous substances. Machine learning algorithms should enable the differentiation between healthy, pre-malignant, and malignant tissue, and more precise delineation of cancer margins for image-guided biopsy and surgery, with a solution for the fast processing.

The third challenge lies in the establishment of a large spectra database for important molecular biomarkers and all types of tissue, including skin and subcutaneous tissue, ocular tissue, head/brain tissue, epithelial/mucous tissue, breast tissue, cartilage, liver, muscle, aorta, lung, and myocardium. Such a database will make it possible to distinguish not only between oxygenated and deoxygenated blood, but also different tissue types, such as bile duct and the fatty tissue surrounding it [136]. The identification of these imagable biomarkers can also benefit early cancer detection.

Combination with other imaging modalities such as pre-operative Positron Emission Tomography (PET), and intraoperative ultrasound, can leverage the key benefits of each technique individually, overcome the penetration limitation of HSI into biological tissue [147], and broaden the application fields of HSI. In clinical settings, HSI can be easily adapted to conventional diagnostic tools such as endoscope andcolposcope to meet demanding requirements by various medical applications. Multimodal imaging combining reflectance and fluorescence has the potential of revealing more information about tissue under investigation.

The clinical applicability of MHSI is clearly still in its adolescence and requires much more validation before it can be used safely and effectively in clinics. With the

advancement of hardware technologies, image-analysis methods and computational power, we expect that HSI will play an important role for noninvasive disease diagnosis and monitoring, identification and quantitative analysis of cancer biomarkers, image-guided minimum invasive surgery, targeted drug delivery and tracking, and pharmaceutical drug dosage assessment.

CHAPTER 3

NONINVASIVE CANCER DETECTION IN SUBCUTANEOUS CANCER MODELS USING HYPERSPECTRAL IMAGING AND SPECTRAL-SPATIAL CLASSIFICATION

In this chapter, we developed a spectral-spatial classification method to distinguish cancer from normal tissue on hyperspectral images of a subcutaneous cancer model. Specifically, we first represented the neighborhood of each pixel as a third-order tensor to preserve the local spectral-spatial structure of a hypercube, and then employed Tucker tensor decomposition to extract the spectral-spatial feature for each pixel. Next, we fed the the extracted tensor features into support vector machines to classify each pixel as cancer or normal tissue. To validate the proposed algorithm, we acquired hyperspectral reflectance images from 450 nm to 900 nm with a 2-nm increment from tumor-bearing mice. The preliminary study demonstrated the feasibility of using HSI for noninvasive detection of tumors in a preclinical animal model.

3.1 Introduction

Survival and life quality of the patients correlate directly to the size of the primary tumor at first diagnosis, therefore, early detection of malignant lesions could improve the chance of survival and lower the rate of recurrence [148]. Suspicious lesions found through standard screening methods should be biopsied for histopathological assessment to make definitive diagnosis [10]. Due to the heterogeneous morphology and visual appearance of the lesions, biopsy diagnosis may not be representative of the highest pathological grade of a tumor due to the small sampling area [11]. After biopsies, tissue

samples are sectioned and stained. Pathologists then examine the specimens under microscopes and make judgments based on observations of cell morphology and colors of different tissue components. This biopsy procedure is time-consuming and invasive. In addition, the interpretation of the histological slides is subjective and can be inconsistent due to intra-observer and inter-observer variation [10] [12].

Hyperspectral imaging (HSI) has the potential to improve cancer diagnostics, decrease the use of invasive biopsies, and reduce patient discomfort associated with traditional procedures [37]. Hyperspectral images, known as a hypercube, are three-dimensional (3D) dataset (x, y, λ), comprised of two spatial dimensions (x, y) and one wavelength dimension (λ). As illustrated in Figure 3, each plane of the hypercube represents a grayscale image at a particular wavelength, and intensities over all the spectral bands form a spectral signature for each pixel of the hypercube. The spectral signature of each pixel has hundreds of contiguous bands covering the ultraviolet, visible, and near-infrared wavelength range.

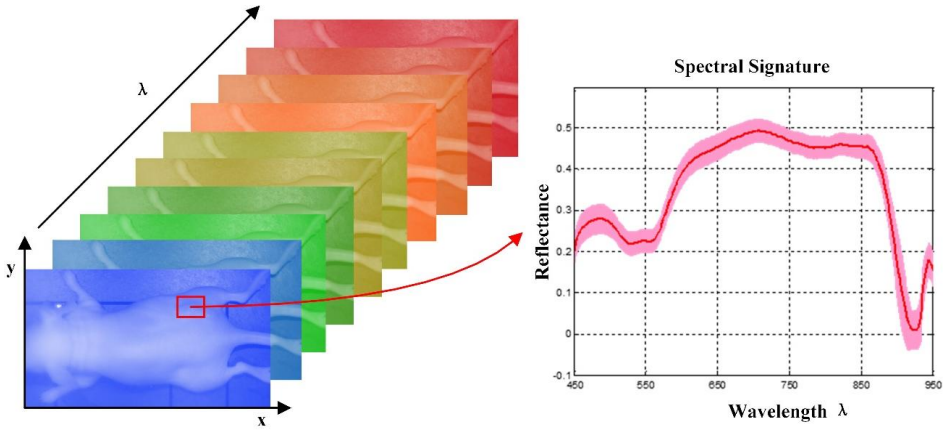


Figure 3: The data structure of a hypercube. The red solid line represents the average reflectance spectrum of the rectangular region of the tumor tissue in the mouse, and the color region around the solid line represents the standard deviation of the spectra in the same region.

The HSI system employed in this study, which operates in the wavelength range of 450-900 nm with a 2-nm interval, with an image size of 1040 × 1392 pixels in the spatial dimension, will generate 1.45 million spectra in one hypercube, each with 226 data points. Machine learning techniques can be applied to mine the vast amounts of data generated in HSI experiments to extract useful diagnostic information and to classify each pixel into cancerous or healthy tissue type.

Traditional classification methods for hyperspectral imaging mainly consist of spectral classification and spatial classification. Spectral classification methods only relied on the spectral signature of each pixel in hyperspectral images. For example, Liu *et al.* [56] proposed a classification method based on the sparse representation of the reflectance spectra for tongue tumor detection from human tongue hyperspectral data of 81 channels from 600 to 1000 nm, and achieved an accuracy of 96.5%. Based on the spectral characteristics of tissues, our group used hyperspectral data (450-950 nm, 251 channels) and a support vector machine (SVM)-based classifier for prostate tumor detection [78]. Akbari *et al.* used hyperspectral data (1,000 – 2,500 nm) for the detection of gastric cancer [81]. Spatial classification methods only employed the spatial information for labeling cancerous tissue samples. Masood *et al.* used hyperspectral images of colon biopsy samples and a single band to classify the sample as normal or cancerous tissue based on the texture feature-circular local binary pattern [83] [112] and wavelet texture features [111]. The spectral methods utilized the spectral signature of individual pixels without considering the spatial relationship of neighboring pixels. Spatial-based methods were limited to one spectral band without fully exploiting the spectral information in hyperspectral data. Therefore, knowledge of how to incorporate spatial and spectral information in a low dimensional space is critical for improving the interpretation and classification of hyperspectral data.

In view of the wealth information available from hyperspectral imaging and the biochemical complexity of tumors, we propose a spectral-spatial classification method

that utilizes the entire spectra at each pixel as well as the information from its neighborhood to differentiate between cancerous and normal tissue. Feature extraction and dimension reduction is an important step to extract the most relevant information from the original data and represent that information in a low dimensional space. Dimension reduction methods, such as principal component analysis (PCA) [149], independent component analysis (ICA), maximum noise fraction (MNF), and sparse matrix transform [150] [151] require spatial rearrangement by vectorizing the 3D hypercube into two-way data, leading to a loss of spatial information. We propose to preserve the local spectral-spatial structure of the hypercube by tensor computation and modeling. The tensor provides a natural representation for hyperspectral data. In the remote sensing area, tensor modeling has been increasingly utilized for target detection [152], denoising [153], dimensionality reduction [154] [155], and classification [156-159]. We extract low-dimensional spectral-spatial features by Tucker tensor decomposition [160], and generate probability maps using support vector machine (SVM) to indicate how likely each pixel is cancerous. The classification method is generic, which can be applied to not only hyperspectral images but also to other medical images such as MRI and CT images. In this study, we demonstrate the efficacy of hyperspectral imaging in combination with spectral-spatial classification methods for *in vivo* head and neck cancer detection in an animal model. The experimental design and methods are described in the following sections.

3.2 Materials

3.2.1 Hyperspectral Imaging Instrument

The HSI system used in this experiment has been described in Chapter 2.

3.2.2 Animal Model

In our experiment, a head and neck tumor xenograft model using HNSCC cell line M4E (doubling rate: ~ 36 hours) was adopted. The HNSCC cells (M4E) were maintained as a monolayer culture in Dulbecco's modified Eagle's medium (DMEM)/F12 medium (1:1) supplemented with 10% fetal bovine serum (FBS) [161]. M4Ecells with green fluorescence protein (GFP), which were generated by transfection of pLVTHM vector into M4E cells, were maintained in the same condition as M4E cells. Animal experiments were approved by the Animal Care and Use Committee of Emory University. Female nude mice aged 4-6 weeks were injected with 2 x 10^6 M4E cells with GFP on the back of the animals. Hyperspectral images were obtained about two weeks post cell injection.

3.2.3 Reference Image Acquisition

Prior to the animal image acquisition, white reference image cubes were acquired by placing a standard white reference board in the field of view with an auto-exposure setting. The dark reference cubes were acquired by keeping the camera shutter closed. These reference images were used to calibrate hyperspectral raw data before image analysis [60].

3.2.4 Reflectance Image Acquisition

During the image acquisitions, we first scan the mice using the reflectance mode. Hyperspectral reflectance images were acquired by anesthetizing each mouse with a continuous supply of 2% isoflurane in oxygen. The excitation setting used the interior infrared (800-900 nm) excitation and the white light excitation (450-800 nm). The acquisition wavelength region for reflectance images was set from 450 to 900 nm with a 2-nm increment. The exposure time was set by the auto-exposure configuration. To eliminate the effect of GFP signals in the reflectance images, the emission bands of GFP at 508 nm and 510 nm were removed in the data preprocessing. Hence, the resultant reflectance images contain 224 spectral bands.

3.2.5 Fluorescence Image Acquisition

Hyperspectral fluorescence images were subsequently acquired without moving the mouse. The blue excitation light at 455 nm and blue emission filter at 490 nm were used to generate GFP fluorescence images. The exposure time was also set as auto-exposure. Tumors show green signals in fluorescence images due to the GFP in tumor cells, and their positions are exactly the same as that in reflectance images. Therefore, GFP fluorescence images can be used as the *in vivo* gold standard for the classification evaluation of cancer tissue on hyperspectral imaging data.

3.2.6 Histological Processing

After data acquisitions, mice were sacrificed by cervical dislocation. Tumors were cut horizontally, and were then put into formalin. Histological slides were prepared from the tissue specimens for further analysis. The histological diagnosis results were also used to confirm the cancer diagnosis.

3.3 Methods

In this section, we explain in detail our spectral-spatial classification method. Figure 4 (a) illustrates the traditional pixel-wise spectral method which only utilizes spectral information. Vectorization of the 3D hypercube into 2D matrix causes loss of spatial information. Principal component analysis (PCA) is usually applied to reduce data dimension. Figure 4 (b) represents the flowchart for the proposed spectral-spatial method. After the input hypercube is preprocessed, a spectral-spatial tensor representation, which conserves the 3D hypercube structure, is constructed. Tensor decomposition is then performed to extract important features and reduce dimension. A support vector machine (SVM) classifier is then applied to classify each pixel into cancerous or healthy tissue with probability estimates. Any classifier that can provide cancer probability estimates

with good classification performance can be applied in this step. Finally an active contour based method is performed as a post-processing step to refine the classification results.

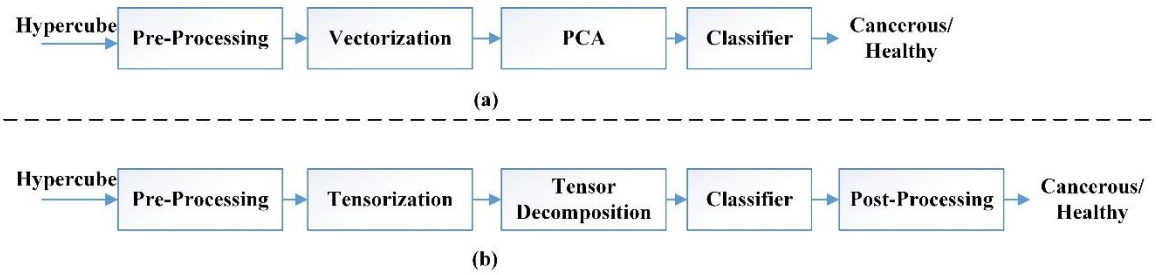


Figure 4: The flowchart of the tissue classification methods. (a) The traditional spectral-based classification method. (b) The proposed spectral-spatial classification method.

3.3.1 Pre-processing

Hyperspectral data preprocessing aims at removing the effects of the imaging system noise and compensating for geometry-related changes in image brightness. It consists of the following four steps:

Step 1: Reflectance Calibration. The purpose of reflectance calibration was to remove the spectral non-uniformity of the illumination device and the influence of the dark current. The raw data can be converted into normalized reflectance using Equation (1).

Step 2: Curvature Correction. In clinical applications, curvature correction is particularly useful when the surface areas of the cancer are raised or depressed with respect to the surrounding tissue. For example, the normal colon tissue surface has numerous folds, and normalization has to be applied to compensate for the difference in the intensity of the light recorded by the camera as a function of tissue geometry [162]. In our experiment, the tumor surface was raised compared to its surrounding normal tissue. So it was desirable to perform curvature correction to compensate for the spectral

variations caused by the elevation of the tumors. The light intensity changes could be viewed as a function of the distance and the angle between the surface and the detector. Two spectra of the same point acquired at two different distances and/or inclinations will have the same shape but vary by a constant [162]. By dividing each individual spectrum by a constant calculated as the total reflectance across the wavelength range removes the distance and angle dependence as well as dependence on an overall magnitude of the spectrum. This normalization step ensures that the variation in spectra curves is only a function of wavelength and therefore the differences between cancerous and normal tissue is not affected by the changing curvature of tumors.

Step 3: Noise Removal. Filters are commonly used for denoising in medical images [163] [164]. After the normalization in Steps 1 and 2, the tissue spectra still presents some noise, which might be due to the breathing of the mice, or small food residuals. Therefore, a median filter is applied to eliminate spectral spikes and to smooth the spectral curves at each pixel, while retaining the variations across different wavelengths.

Step 4: GFP Bands Removal. GFP signal produces a strong contrast between tumor and normal tissue under blue excitation, and may also present a good contrast compared to other spectral bands under white excitation. To eliminate the effect of GFP signals on the cancer detection process, GFP spectral bands, *i.e.* 508 nm and 510 nm in our case, are removed from the image cubes before feature extraction.

3.3.2 Spectral-Spatial Tensor Representation

Tensors are generalizations of matrices and vectors. A first-order tensor is a vector, a second-order tensor is a matrix, and tensors of order three or higher are called higher-order tensors [165]. The order of a tensor is the number of dimensions, also known as modes. An N-way or N^{th}-order tensor $\underline{\mathbf{X}} \in \mathbb{R}^{J_1 \times J_2 \times \cdots \times J_N}$ is represented by a multidimensional array with N indices. In this study, hyperspectral data $\underline{\mathbf{X}}$ is a set of

$J_3 = 224$ images, corresponding to wavelength band from 450 to 900 nm. Each spectral image is composed of $J_1 \times J_2$ pixels, with x_{j_1, j_2, j_3} representing the intensity at pixel (j_1, j_2) in spectral band j_3.

To fully exploit the natural multi-way structure of hyperspectral data, we construct a spectral-spatial tensor representation for each pixel, by dividing hypercube ($J_1 \times J_2 \times J_3$) into overlapping patches of $3 \times 3 \times J_3$ dimension. Hence, each pixel is represented by a third-order tensor $\underline{\mathbf{X}} \in \mathbb{R}^{3 \times 3 \times J_3}$, with two modes representing spatial information, and the third mode for spectral band. $J_3 = \lambda$ represents the number of spectral bands. Figure 5 illustrates the spectral-spatial tensor representation of the hypercube. With the spectral-spatial tensor representation, spectral continuity is represented as the third tensor dimension, while spatial information is included as row-column correspondence in the mathematical structure.

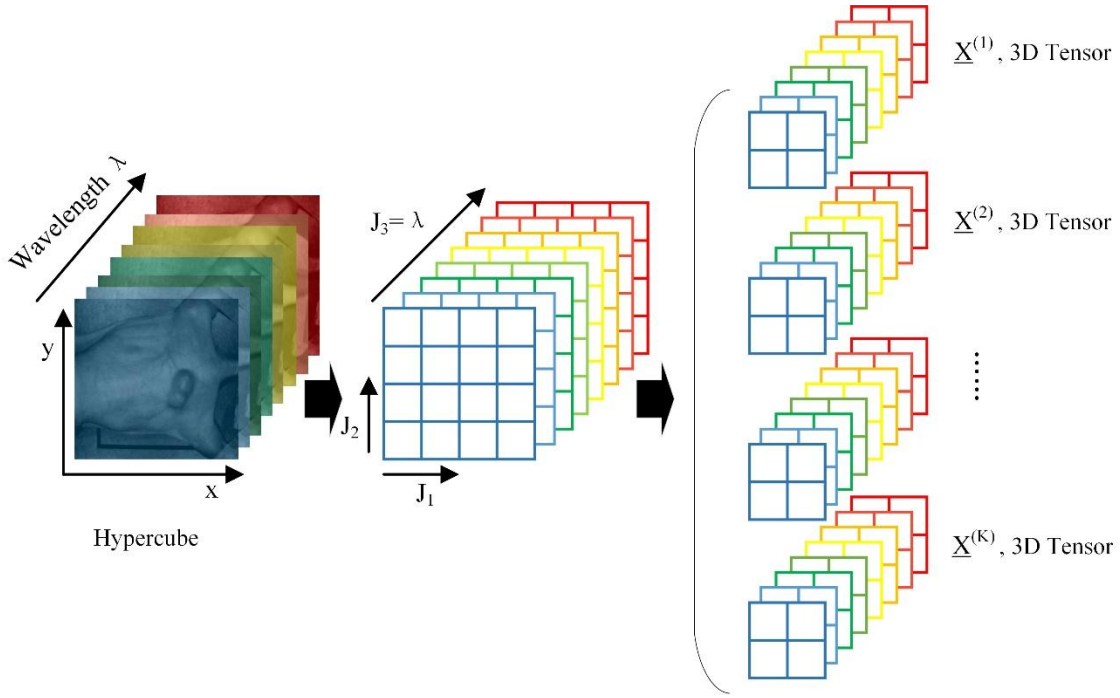


Figure 5: Spectral-spatial tensor representation of a hypercube. The image stack on the left is the hypercube of a tumor-bearing mouse. The image stack in the middle shows that a hypercube ($J_1 \times J_2 \times J_3$) can be divided into small patches. The image stack on the right shows that each pixel inside a hypercube can be represented by a small patch centered at that pixel. This patch containing information from both the pixel and its neighborhood can be represented in a mathematical form as a 3D tensor.

3.3.3 Feature Extraction and Dimension Reduction

Tensor decompositions are important tools for feature extraction and dimension reduction by capturing the multi-aspect structures of the large scale high dimensional data, with applications in image and signal analysis, neuroscience, and chemometrics. Tucker tensor decomposition is a basic model for high dimensional tensors which allows effective feature extraction and dimension reduction. An N-way Tucker tensor $\underline{\mathbf{X}}^k \in \mathbb{R}^{J_1 \times J_2 \times \cdots \times J_N}$ can be decomposed into a core tensor $\underline{\mathbf{G}}^k \in \mathbb{R}^{R_1 \times R_2 \times \cdots \times R_N}$ multiplied or transformed by a set of component matrices $\mathbf{A}^{(n)} = \left[\mathbf{a}_1^{(n)}, \mathbf{a}_2^{(n)}, \cdots, \mathbf{a}_{R_n}^{(n)}\right] \in \mathbb{R}^{J_n \times R_n}$ (n=1,2,…N) [160]:

$$\underline{X}^k = \sum_{j_1=1}^{R_1} \sum_{j_2=1}^{R_2} \cdots \sum_{j_N=1}^{R_N} g_{j_1 j_2 \cdots j_n} \left(a_{j_1}^{(1)} \circ a_{j_2}^{(2)} \circ \cdots a_{j_N}^{(N)} \right) + \underline{\mathcal{E}}^k \quad (2)$$

$$= \underline{\mathbf{G}}^k \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \cdots \times_N \mathbf{A}^{(N)} + \underline{\mathcal{E}}^k$$

$$= \underline{\widehat{\mathbf{X}}}^k + \underline{\mathcal{E}}^k$$

Where symbol "∘" represents outer product, $\underline{\mathbf{G}}^k \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \cdots \times_N \mathbf{A}^{(N)}$ denotes the multiplication in all possible modes of a tensor $\underline{\mathbf{G}}^k$ and a set of matrices $\mathbf{A}^{(n)}$. In practice, it is common that the core tensor $\underline{\mathbf{G}}^k$ is smaller than the original tensor $\underline{X}^k$, i.e. $R_N < J_N$. Decomposition of tensor $\underline{X}^k$ can be seen as a composition of directional bases $\mathbf{A}^{(1)}, \mathbf{A}^{(2)} \cdots \mathbf{A}^{(N)}$ in modes 1,2, … , N, connected through a set of weights contained in $\underline{\mathbf{G}}^k$. The elements in the core tensor $\underline{\mathbf{G}}^k$ represents the features of the sample $\underline{\mathbf{X}}^k$ in the subspace spanned by $\mathbf{A}^{(n)}$. Hence, the extracted features are usually in a lower dimension than the original data tensor $\underline{\mathbf{X}}^k$.

We assume that the basis matrices $\mathbf{A}^{(n)} = \left[\mathbf{a}_1^{(n)}, \mathbf{a}_2^{(n)}, \cdots, \mathbf{a}_{R_n}^{(n)} \right] \in \mathbb{R}^{J_n \times R_n}$ (n=1,2,…N) are common factors for all data tensors. $\underline{\widehat{\mathbf{X}}}^k$ is an approximation of $\underline{\mathbf{X}}^k$, and $\underline{\mathcal{E}}^k$ presents the approximation error. Figure 6 illustrates the Tucker decomposition of a three-way tensor. To compute the basis matrices $\mathbf{A}^{(n)}$ and the core tensor $\underline{\mathbf{G}}^k$, we concatenate all individual tensors $\underline{\mathbf{X}}^k$ into one N+1 order training tensor $\underline{\mathbf{X}} = \text{cat}\left(\underline{\mathbf{X}}^1, \underline{\mathbf{X}}^2, \cdots, \underline{\mathbf{X}}^K, \text{N+1}\right) \in \mathbb{R}^{J_1 \times J_2 \times \cdots \times J_N \times J_{N+1}}$ with $J_{N+1} = K$ and perform Tucker-N decomposition [160]. The sample tensors $\underline{\mathbf{X}}^k$ can be obtained from the concatenated tensor by fixing the (N+1)-th index at a value k and the individual features can be extracted from the core tensor $\underline{\mathbf{G}}^k$ by fixing the (N+1)-th index as k. In the case of hyperspectral data, the core tensor connects two spatial modes with one spectral mode of the hypercube. Hence, the extracted features simultaneously contain the spectral-spatial profile of tissues in the hyperspectral data $\underline{\mathbf{X}}^k$.

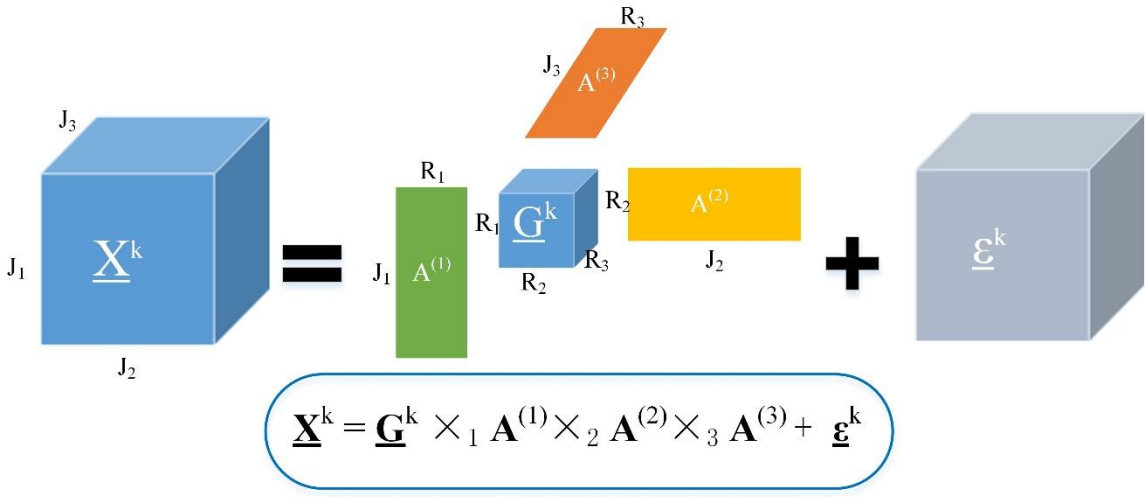


Figure 6: Tucker decomposition of a three-way tensor $\underline{\mathbf{X}}^k$. Decomposition of tensor $\underline{\mathbf{X}}^k$ can be seen as a multiplication in all possible modes of a core tensor $\underline{\mathbf{G}}^k$and a set of basis matrices $\mathbf{A}^{(n)}$.

In general, Tucker decomposition is not unique [165]. Constraints such as orthogonality, sparsity, and nonnegativity are commonly imposed on the component matrices and the core tensor of the Tucker decomposition, in order to obtain meaningful and unique representation [166]. To solve the Tucker tensor decomposition problem, we applied the higher order discriminant analysis (HODA) with orthogonality constraints on basis factors [160], which is a generalization of linear discriminant analysis (LDA) for multi-way data. HODA aims to find discriminant orthogonal bases to project the training features $\underline{\mathbf{G}}^k$ onto the discriminant subspaces. Optimal orthogonal basis factors $\mathbf{A}^{(n)}$ can be found by maximizing the Fisher ratio between the core tensors $\underline{\mathbf{G}}^k$:

$$\max_{\mathbf{A}^{(1)},\mathbf{A}^{(2)},\ldots\mathbf{A}^{(N)}} \frac{\sum_{c=1}^{C} K_c \left\| \bar{\underline{\mathbf{G}}}^c - \bar{\bar{\underline{\mathbf{G}}}} \right\|_F^2}{\sum_{k}^{K} \left\| \underline{\mathbf{G}}^k - \bar{\underline{\mathbf{G}}}^{c_k} \right\|_F^2} \tag{3}$$

where $\bar{\underline{\mathbf{G}}}^c$ is the mean tensor of the c-th class consisting of K_c training samples, and $\bar{\bar{\underline{\mathbf{G}}}}$ is the mean tensor of the whole training features. c_k denotes the class label of the k-th

training sample $\underline{\mathbf{X}}^{k}$. The details for solving the above optimization problem can be find in [160].

The dimension of the extracted feature is $L = R_1 \times R_2 \times \cdots \times R_N$, which is dependent on the dimension of the basis factors $\mathbf{A}^{(n)} \in \mathbb{R}^{J_n \times R_n}$. R_n can be determined by the number of dominant eigenvalues of the contracted product $\mathbf{X}_n \mathbf{X}_n{}^T \in J_n \times J_n$, $\mathbf{X}_n \mathbf{X}_n{}^T = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\boldsymbol{T}}$, where $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots \lambda_{J_n})$ and $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{R_n} \geq \cdots \geq \lambda_{J_n}$ are eigenvalues. The optimal dimension R_n of the core tensor can be found by setting a threshold fitness θ and optimize the following problem:

$$\arg\min_{R_n} \frac{\sum_{i=1}^{R_n} \lambda_i}{\sum_{i=1}^{J_n} \lambda_i} > \theta \tag{4}$$

In our experiment, we set $\theta = 99$, which means the factors should explain the whole training data in at least 99% of the cases.

The approach of feature extraction for both training and testing data is illustrated in Figure 7. After spectral-spatial representation, the training data is constructed by concatenating K sample patches as a 4-D tensor of size $J_1 \times J_2 \times J_3 \times K$, and the testing data is formed in the same manner. Here, we choose grid size of $J_1 \times J_2 = 3 \times 3$ and the wavelength dimension of $J_3 = 224$. We first perform the third-order orthogonal Tucker tensor decomposition along the mode-4 on the training data using HODA. After the Tucker decomposition, the core tensor $\underline{G} \in \mathbb{R}^{R_1 \times R_2 \times R_3}$, which expresses the interaction among basis components, is vectorized into a feature vector with a length of $R_1 \times R_2 \times R_3$ as the training feature. The dimension of the tensor feature can be much less than that of the original pixel-based feature. Therefore, dimension reduction can be achieved by projecting the original tensors X to the core tensors G with proper dimensions for R_1, R_2, and R_3. To extract features from testing data, the basis matrices $A^{(n)}$ found from training data is used to calculate the core tensor and the corresponding core tensor is then converted into a testing feature vector. If the feature dimension is still high after the

feature extraction step, feature ranking or feature selection can be applied to further reduce the feature dimension. Finally, the extracted testing features are compared with the training features using SVM classifiers and the probability map for cancerous tissue is generated.

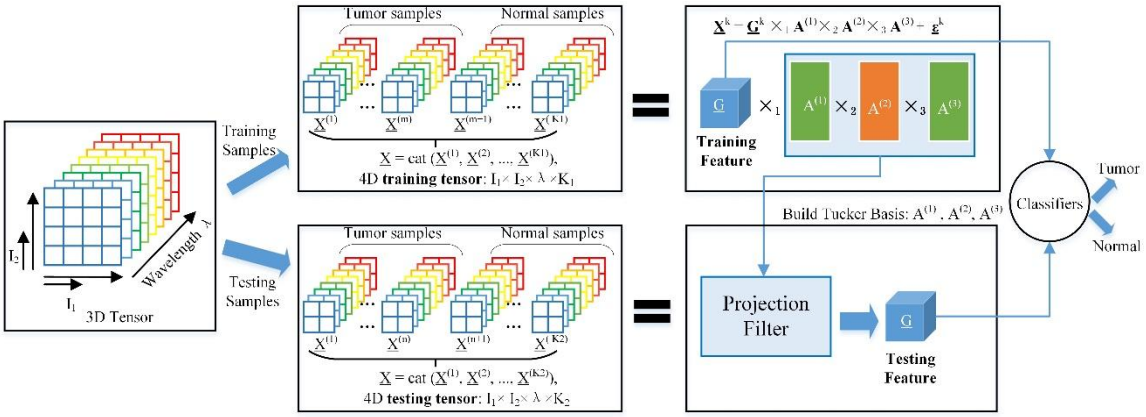


Figure 7: Feature extraction using Tucker tensor decomposition.

3.3.4 Classification

In this study, 12 hypercubes from 12 mice with head and neck cancers are scanned and used for the hyperspectral image analysis. We choose support vector machine as the classifier and the Gaussian radial basis function (RBF) as the kernel function [167]. Nested cross validations are used to perform model selection and evaluation. We perform leave-one-out outer cross validation, and three-fold inner cross validation (CV). A grid search is performed in the inner cross validation on the training data to select the optimal value for parameter C and g over the range of $\log_2 C = \{-1, 0, \cdots, 4\}$ and $\log_2 g = \{-1, 0, \cdots, 4\}$. Then a new SVM model is trained with the optimal parameters on eleven mouse data; and the performance of that model is tested on the remaining mouse.

3.3.5 Post-Processing

After obtaining the probability maps of each tumor image, we proceed to use active contours [168] [169] to refine the classification results. Chan-Vese active contour [170] is a region-based segmentation method, which can be used to segment a vector-valued image such as RGB images. Standard L2 norm is used to compare the image intensity with the mean intensity of the region inside and outside the curve. However, if the image contains artifacts, the L2 norm may not work well. Therefore we modified the Chan-Vese active contour with the L1 norm which compares the image intensity with the median intensity of the region inside and outside the curve. This modification makes the active contour more robust to noises. In this section, we will first introduce the mathematic formulation of the modified L1 norm active contour method.

The energy function with the L1 norm is defined as follows:

$$\mathrm{E(u,v,C)} = \iint_{\Omega} \frac{1}{N}\sum_{i=1}^{N}|I_i - u_i|\,dxdy + \iint_{\Omega^c} \frac{1}{N}\sum_{i=1}^{N}|I_i - v_i|\,dxdy + \lambda \int_C ds \qquad (5)$$

where C stands for the curve, Ω stands for the area inside C, and Ω^c stands for the area outside C. The last term penalizes the "shape" of the curve to avoid complicated curves. N is the total number of image bands.

Given the curve C, we want to find the optimal values of u_i and v_i. By setting $\frac{\delta \mathrm{E(u,v,c)}}{\delta u_i} = 0$ and $\frac{\delta \mathrm{E(u,v,c)}}{\delta v_i} = 0$, we obtain:

$$\frac{\delta \mathrm{E(u,v,c)}}{\delta u_i} = \frac{\delta}{\delta u_i}\iint_{\Omega} \frac{1}{N}\sum_{i=1}^{N}|I_i - u_i|\,dxdy$$

$$= \frac{\delta}{\delta u_i}\iint_{\Omega} \frac{1}{N}\sum_{i=1}^{N}\sqrt[2]{(I_i - u_i)^2}\,dxdy$$

$$= -\frac{1}{N}\iint_{\Omega} \frac{I_i - u_i}{|I_i - u_i|}dxdy = 0$$

$$\Rightarrow \iint_{\Omega} \frac{I_i - u_i}{|I_i - u_i|}dxdy = 0 \qquad (6)$$

As we know that $\frac{I-u_i}{|I-u_i|}$ is either 1 or -1, to make the integration of $\frac{I-u_i}{|I-u_i|}$ inside the curve zero, half of the values should be 1, and the other half of the values should be -1, therefore, the optimal value for u_i is the median intensity of the ith image band inside the curve C. Similarly, we know that the optimal value for v_i is the median intensity of the i^{th} image band outside the curve C.

It is expected that the modified active contour with L1 norm applied on the RGB probability maps of tumors will further boost the classification performance.

3.3.6 Comparison with the Spectral-based Classification Method

To compare the proposed spectral-spatial classification method with the spectral-based method for classifying cancerous and normal tissue, we implement the traditional pixel-wise spectral method as illustrated in Figure 4 (a).

Pixel-Wise Spectral-based Method: If no dimension reduction technique is used, then the normalized reflectance spectrum of each pixel with 224 dimensions is directly used as the spectral feature. This method is time-consuming due to the high feature dimension.

Principle Component Analysis (PCA) based Spectral Method: Considering the high dimensions (over 200) of reflectance spectra, PCA is usually applied to reduce the dimensionality. First, the hypercube ($x \times y \times \lambda$) is rearranged into a 2D spectral matrix of dimension $M \times \lambda$, where $M = x \times y$ is the total number of pixels, and λ is the total number of wavelength used. So each row represents the reflectance values from all the bands at one pixel. Then the matrix is centered by subtracting the mean values of each column. Afterwards, PCA was performed to calculate the eigenvalues and eigenvectors. Finally, the original hypercube was approximated by the inverse principle component transformation, with the first few bands containing the majority of the variation residing in the original hypercube [171].

3.3.7 Performance Evaluation Metrics

Accuracy, sensitivity, and specificity are commonly used performance metrics for a binary classification task [172] [173]. In this study, accuracy is calculated as a ratio of the number of correctly labeled pixels to the total number of pixels in a test image. Sensitivity measures the proportion of actual cancerous pixels ("positives") which are correctly identified as such in a test image, while specificity measures the proportion of healthy pixels ("negatives") which are correctly classified as such in a test image. F-score is the harmonic mean of precision (the proportion of correct true positives to all predicted positives) and sensitivity. Table 1 shows the confusion matrix, which contains information about actual and predicted classification results performed by a classifier.

Table 1: Confusion Matrix

		Predicted Results	
		Negative (healthy)	Positive (cancerous)
Gold Standard	Negative (healthy)	True Negative (TN)	False Positive (FP)
	Positive (cancerous)	False Negative (FN)	True Positive (TP)

The definitions of accuracy, precision, sensitivity, specificity, and F-score are defined below:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN};\ \text{Sensitivity} = \frac{TP}{TP + FN};\ \text{Specificity} = \frac{TN}{TN + FP}$$

$$\text{Precision} = \frac{TP}{TP + FP};\ F = 2 \times \frac{\text{precision} \times \text{sensitivity}}{\text{precision} + \text{sensitivity}}$$

3.4 Results

To evaluate the proposed tumor detection algorithm, we scanned 12 GFP tumor-bearing mice approximately two weeks post tumor cell injection and distinguished between cancerous and normal tissue based on their spectral differences in this study.

3.4.1 Data Normalization

To visualize the spectral variation arising from tumor curvature, a region of interest covering the tumor of a mouse was selected on the reflectance composite RGB color image (shown in Figure 8 (a) and (d)), and the average spectra for a selection of regions at various locations (denoted by squares in Figure 8 (a) and (d)), were obtained by averaging pixel spectra from a square area of 10×10 pixels from those regions. It is obvious that the curvature of the tumor surface causes a scaling difference in the spectra: the spectrum from the center of the tumor exhibited higher reflectance intensity than the spectra from the side of the tumor and the surrounding normal tissue. This is mainly caused by the relative difference in the path length from different points of the curved tumor surface to the detector. So it is desirable to minimize the spectral variability caused by tumor curvature.

Figure 8 (b) and (e) shows the spectral variation along horizontal direction and vertical direction respectively. Figure 8 (c) and (f) shows the spectral curve of pixels along horizontal and vertical directions after data normalization. The color bar varying from green to black in Figure 8 (b) and (c) represents the location from left to right, while the color bar in (e) and (f) denotes the location from top to bottom. It can be seen that the spectral variance is greatly reduced and the spectral curve is smoothed after applying the pre-processing procedure. It is reasonable to assume that after pre-processing, the spectral variance arises from the actual difference between normal and cancerous tissue.

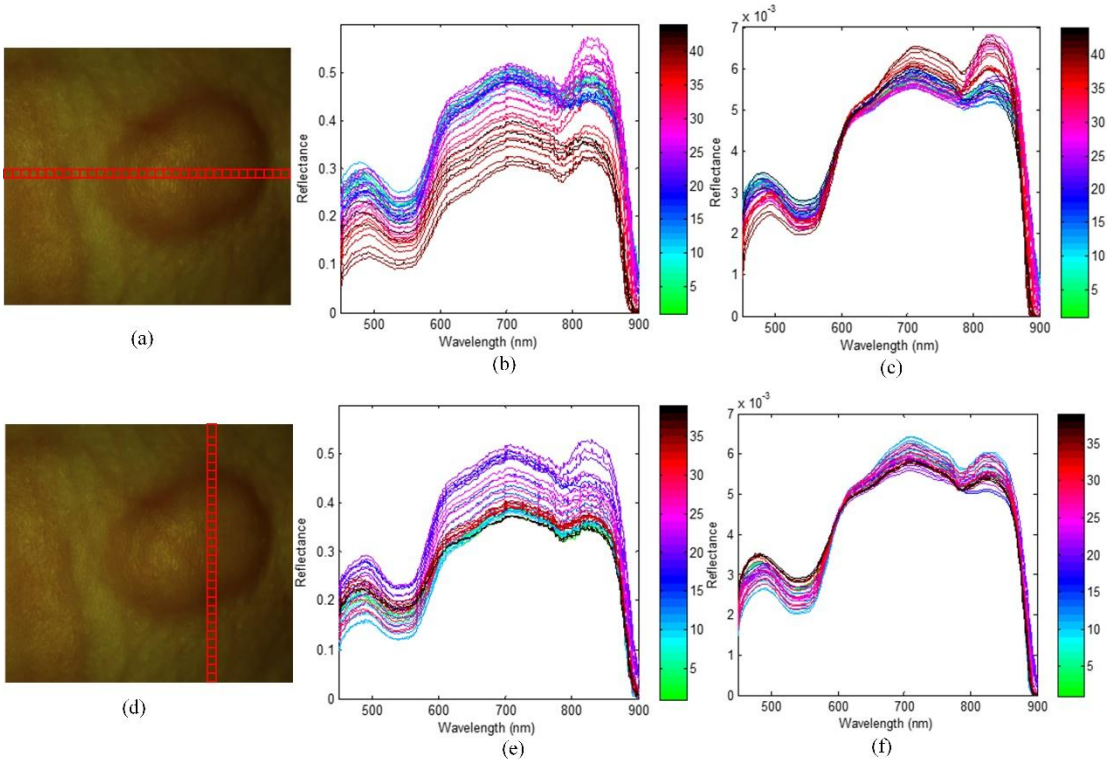


Figure 8: Effects of the pre-processing on spectra as selected from different regions of mouse image. (a) and (d) are the same ROI covering the tumor area; the horizontal and vertical locations are composed of square areas of 10×10 pixels. (b) and (c) show the average spectra of each square from left to right before and after pre-processing. (e) and (f) show the average spectra of each square from top to bottom before and after pre-processing.

3.4.2 Vascularity Visualization

In this study, we acquired hyperspectral reflectance images from both tumors without GFP and tumors with GFP. Figure 9 and Figure 10 show the *in vivo* hyperspectral reflectance images of a tumor without GFP and a tumor GFP, respectively.

To visualize the hyperspectral dataset, RGB composite images were generated as shown in Figure 9 (b) and Figure 10 (b), and the individual image bands at different wavelengths are shown in Figure 9 (a) and Figure 10 (a). In both cases, vascularity patterns can be clearly visualized at different wavelengths. The skin of nude mice, which were covered the tumor, is less than 1 mm thick, so we can visualize tumor vessels

through the intact skin even at 450 nm wavelength. It can be seen that the vascular structure became obscured at higher wavelength, which indicates that light at lower wavelength is more sensitive to superficial information of tissue and that light at higher wavelength carries information from deeper tissue due to deeper penetration into the tissue.

Wavelengths of 508 nm and 510 nm are emission peaks of GFP under blue excitation, and image bands at these two wavelengths for both GFP-tumor and non-GFP-tumor were shown in Figure 9 (a) and Figure 10 (a). It is worth noting that images obtained at GFP bands did not show higher contrast compared to images at other wavelength bands under white excitation. Therefore, it makes sense to assume that the spectral contrast between cancerous and healthy tissue are not caused by the GFP signals.

Figure 9 (c) and Figure 10 (c) show the average reflectance spectra of vessels inside tumor regions with the red solid line, the average reflectance spectra of non-vessel tumor regions with the blue dotted line, and the average of randomly selected normal regions around tumors with the green dash-dot line. The reflectance spectra of both tumors show a dip at around 540-580 nm, which coincides with hemoglobin's absorption peaks. Vessel spectra in the tumor region with wavelengths below 600 nm and above 870 nm exhibit lower reflectance than non-vessel tissue, which is consistent with the higher amount of hemoglobin in vessels. In addition, the reflectance spectra of the tumor region is lower than that of the normal region below 600 nm and above 870 nm, which also indicates that tumors have a higher amount of hemoglobin compared to normal tissue.

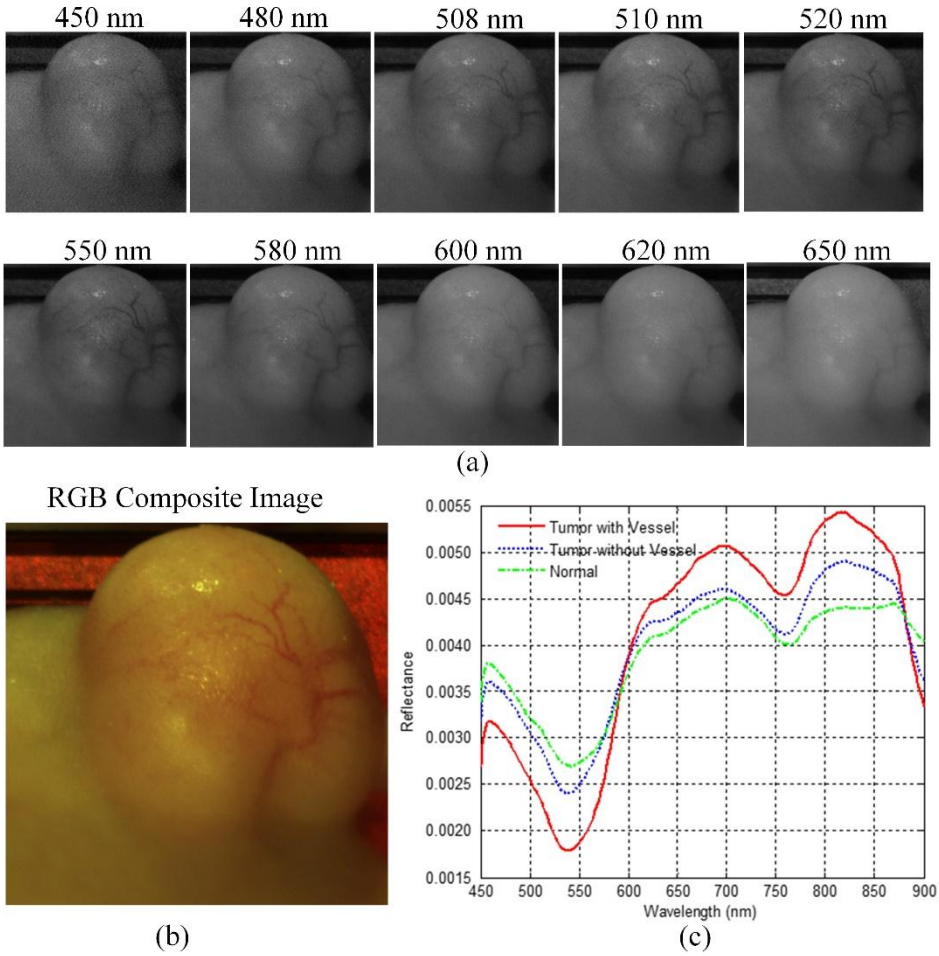


Figure 9: *In vivo* hyperspectral reflectance imaging of a tumor without GFP. (a) Reflectance images at different wavelength bands. (b) A RGB composite image generated from the tumor hypercube. (c) Red solid line: the average reflectance spectra of the vessels inside the tumor region; Blue dotted line: the average reflectance spectra of randomly selected non-vessel tumor regions; Green dash-dot line: the average of randomly selected normal regions around the tumor.

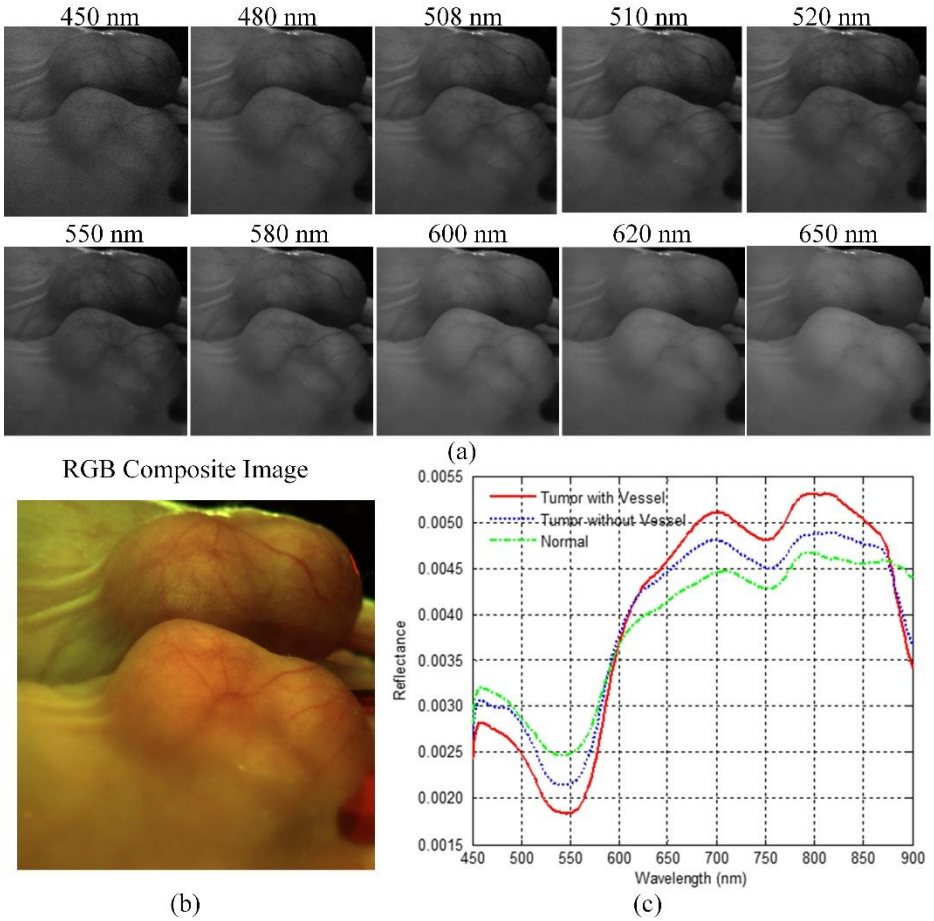


Figure 10: *In vivo* hyperspectral reflectance imaging of a tumor with GFP. A mirror is used to aid in capturing the entire tumor during imaging. (a) Reflectance images at different wavelength bands. 508 nm and 510 nm are the emission peaks for GFP under blue excitation. (b) A RGB composite image generated from the tumor hypercube. (c) Red solid line: the average reflectance spectra of the vessels inside tumor region; Blue dotted line: the average reflectance spectra of randomly selected non-vessel tumor regions; Green dash-dot line: the average of randomly selected normal regions around tumors. (Note: pixels are only selected from the tumor region that is not in the mirror).

3.4.3 Spectra Analysis

Figure 11 shows the RGB composite images of hyperspectral reflectance images for all the twelve mice used for the evaluation of the spectral-spatial classification method. The numbers on the left corners of the mouse images denote the ID on the ear tags of the mice. As can be seen, tumor volumes varied from 92 mm^3 to about 200 mm^3

at the time of imaging. Figure 12 shows the RGB composite images of hyperspectral fluorescence images for tumors with green fluorescence protein (GFP). In both reflectance and fluorescence images, vascular patterns can also be visualized in some tumors, such as # 885 and # 889. Fluorescence images showed vascular beds more clearly and the tumor regions appeared green due to GFP emission peaks at 508 nm and 510 nm.

Although the current gold standard for cancer diagnosis remains histological assessment of hematoxylin and eosin (H&E) stained tissue, the *ex vivo* tissue specimen undergoes deformations, including shrinkage, tearing, and distortion, which makes it difficult to align the *ex vivo* gold standard with *in vivo* tumor tissue. However, the *in vivo* GFP images provided a much better alignment with hyperspectral reflectance images since they were acquired *in vivo* immediately after the acquisition of reflectance images for each mouse, and the tumor and surrounding normal tissue exhibited high contrast in GFP images. In this study, tumor regions were identified manually on the GFP images, and the classification results were then compared with the manual maps. As shown in Figure 12, the GFP images for the tumors #888, #898, #896, #890, and #891 were not clear due to high tissue autofluorescence caused by the food residuals and other wastes on the skin. For these images, we first performed spectral unmixing using the commercial Maestro software to better separate the tumors and surrounding tissues, and then manually segmented the tumors. Since human tissue doesn't contain GFP naturally, registration methods are desirable to align the *in vivo* hyperspectral images with *ex vivo* histological images as discussed in [149] in order to move forward to future human studies.

Figure 13 shows the average reflectance spectra and standard deviations of the cancerous tissue regions from 450 nm to 900 nm in the red solid line, and the average reflectance spectra and standard deviations of the surrounding healthy tissues in the blue dotted line. To make it easier to visualize the two spectra curves, we only showed the standard deviations at the wavelengths from 450 nm to 900 nm with an increment of 50

nm. Due to tumor heterogeneities in morphology, reflectance spectra at different locations varied from each other and deviated from their average reflectance spectra.

Hemoglobin characteristics were shown in these reflectance spectra, which could be indicative of cancer formation. It was found that the normalized reflectance intensity of tumor was lower below 600 nm for all these mice and lower above 870 nm for some of the mice such as #894 and #885. This indicated that the amount of hemoglobin was higher in tumors than normal tissues, which may be due to the angiogenesis during tumor formation.

We have demonstrated that the vascularity patterns can be visualized from hyperspectral reflectance imaging, and the amount of hemoglobin and oxygenated hemoglobin also varies between cancerous and healthy tissue. In addition, vascular density in oral cancers has been shown to be an important biomarker for some cancers [174]. These observations confirm that tissue reflectance spectra measured from 450 nm to 900 nm provide valuable information for differentiating between tumor and normal tissue. Therefore, hyperspectral reflectance imaging has the potential to detect tumors noninvasively through intact skin.

It was noted that the reflectance spectra differences between tumor and normal tissue were relatively small in mice #888, #898, #896, #890, and #891 compared to the rest of the mice, which coincided with the lower contrast between tumor and surrounding tissue in fluorescence images of these mice. This might be because the reflected light from the tissue was further randomly scattered by the food and wastes adhered to the tissue surface, which obscured the differences between cancerous and healthy tissue caused by the biochemical and morphological changes during neoplastic changes.

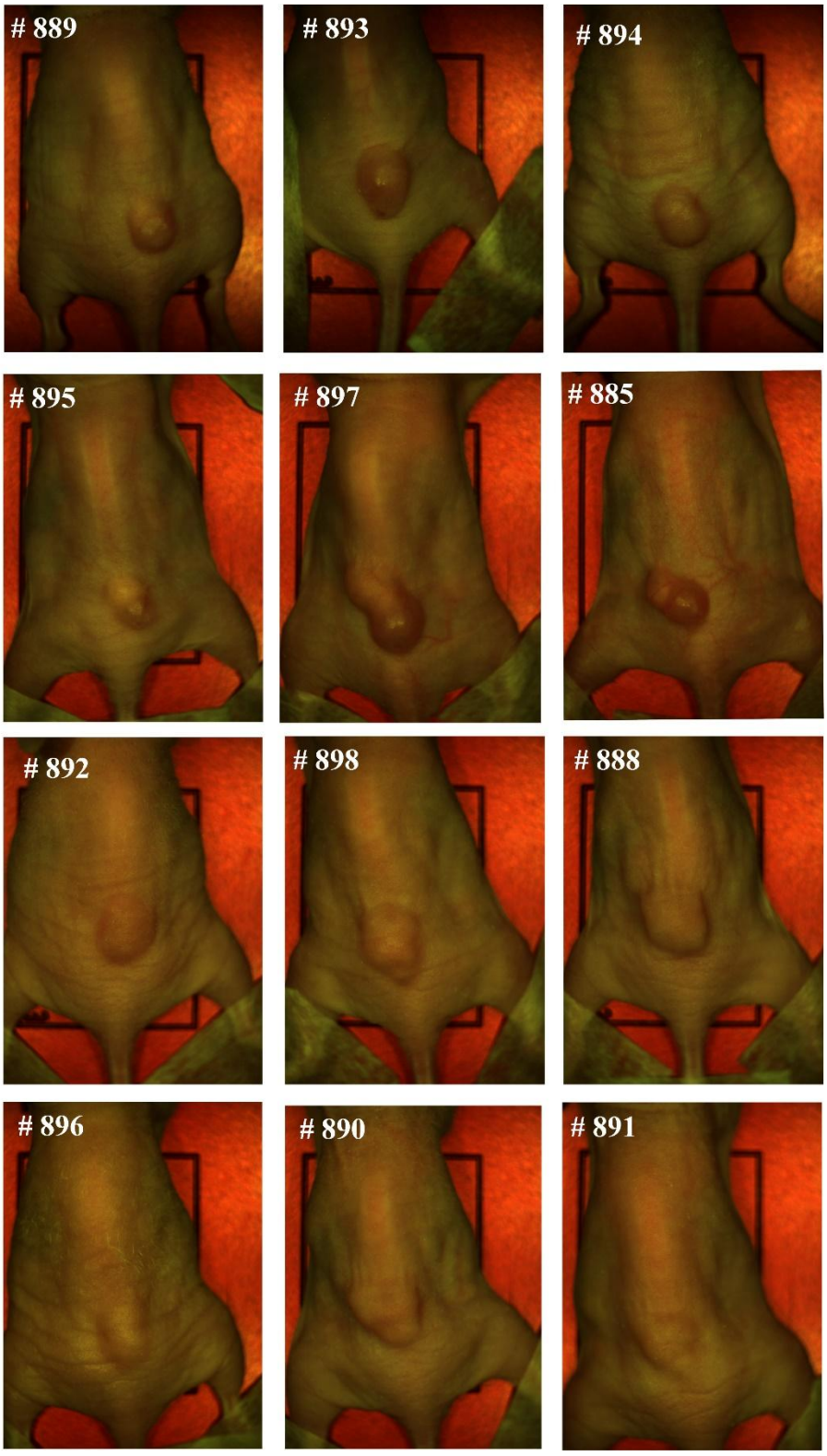


Figure 11: RGB composite images of the reflectance hyperspectral images of the 12 mice used for the evaluation of the spectral-spatial algorithm. The number on the top left corner of each image represents the ID on the ear tag of a mouse.

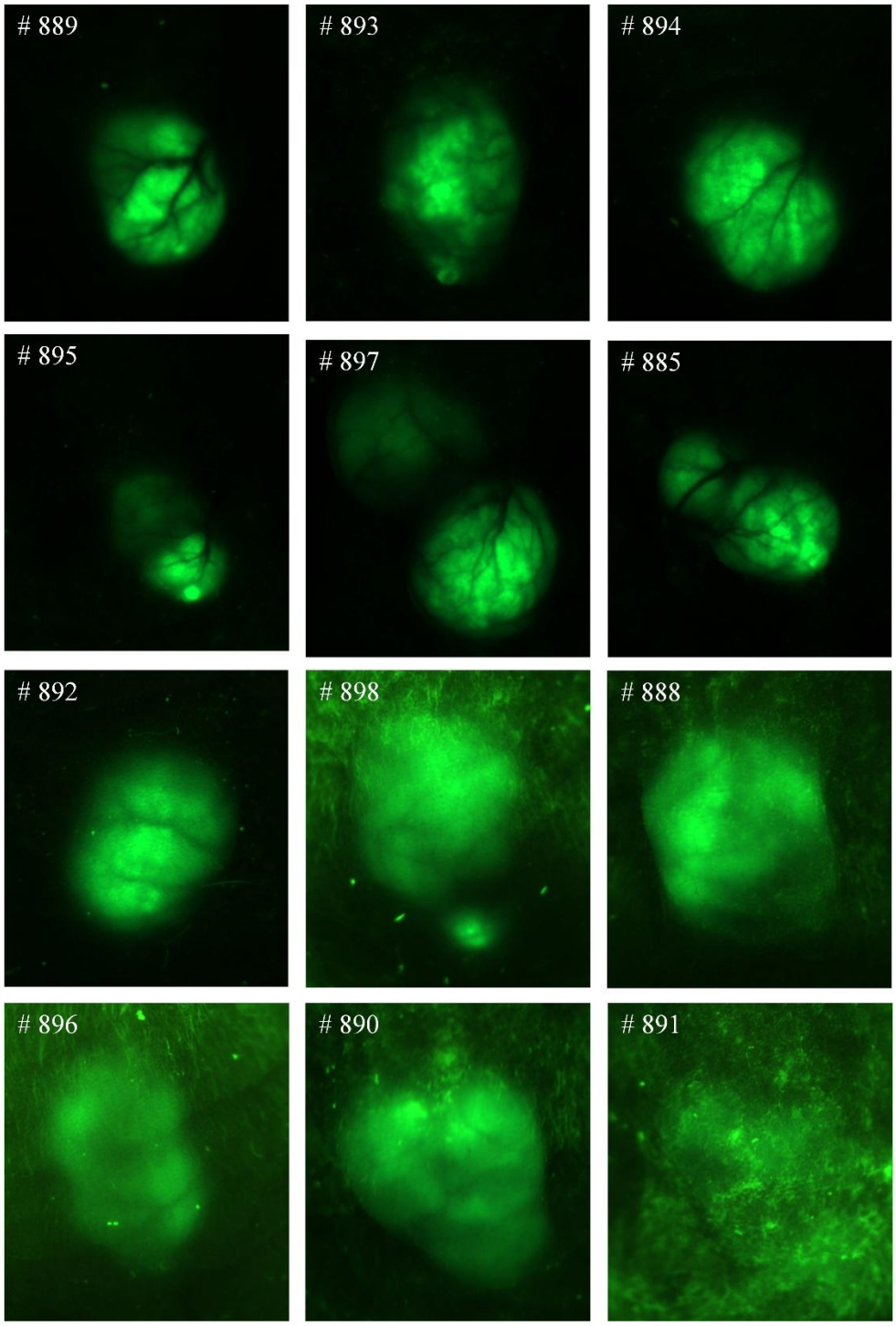


Figure 12: RGB composite images of hyperspectral fluorescence images for the tumors with green fluorescence protein (GFP). Tumors show GFP signals on the images.

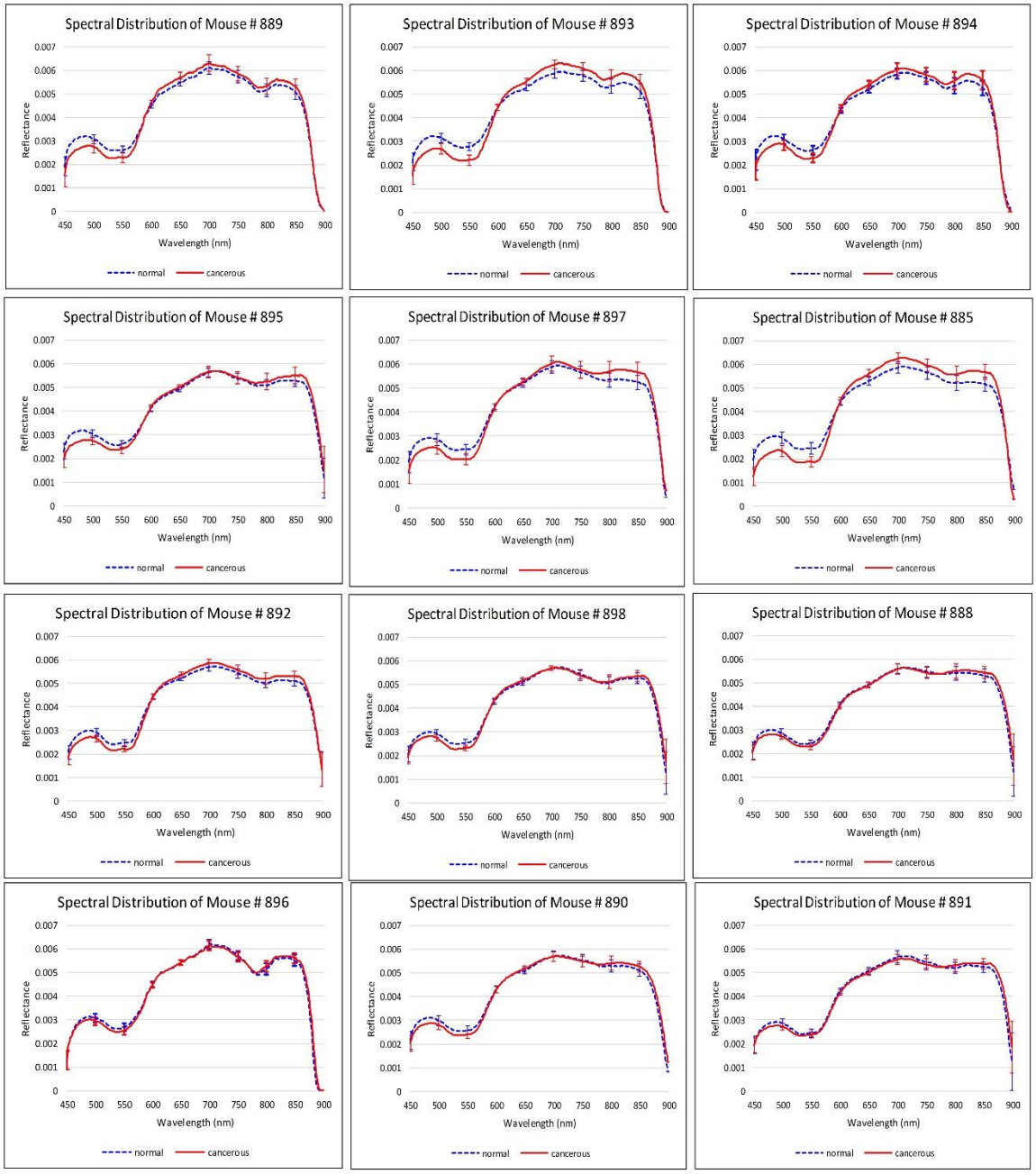


Figure 13: Reflectance spectra of 12 mice. Red solid line: the average spectra of cancerous tissues in each mouse. Blue dotted line: the average spectra of healthy surrounding tissues in each mouse. The error bars in both lines represent the standard deviations of spectra within each region.

3.4.4 Comparison with the Spectral Method

To compare the spectral spatial classification method with the spectral classification methods, we implemented three spectral methods as shown in Figure 14. “Tensor-1” denoted the spectral spatial classification method which represents each pixel with a one-dimensional tensor feature. “Pixel-224” denoted the pixel-wise method which represents each pixel with a vector consisting of 224 reflectance values. “PCA-100” utilized PCA to reduce the pixel-wise feature dimension from 224 to the top 100 most significant features. “PCA-1” utilized PCA to reduce the pixel-wise feature dimension from 224 to the top one significant feature which represents the most variances. For all four methods, KNN classifier is employed to classify the data with leave-one-out cross validation. Figure 14 compared the average and standard deviation of accuracy, sensitivity, specificity, f-score, and precision of the 12 mice for all four methods.

As can be seen from Figure 14, the spectral spatial method with 1D tensor feature outperformed the spectral-based methods. Although the first PCA image band explained about 85% of the variance in the feature vector, the classification accuracy of the PCA-1 method only achieved 54%. More than 100 features were required in order to obtain an accuracy of above 69% with the PCA dimension reduction method. While the top tensor feature alone achieved an accuracy of 80%, which exhibited a strong discriminatory ability for differentiating tumors from normal tissue, it was found that the feature dimension and classification time was significantly reduced without sacrificing the accuracy with the tensor-based spectral spatial method, while higher feature dimension and longer classification time was needed in order to achieve comparable accuracy with the PCA-based spectral method.

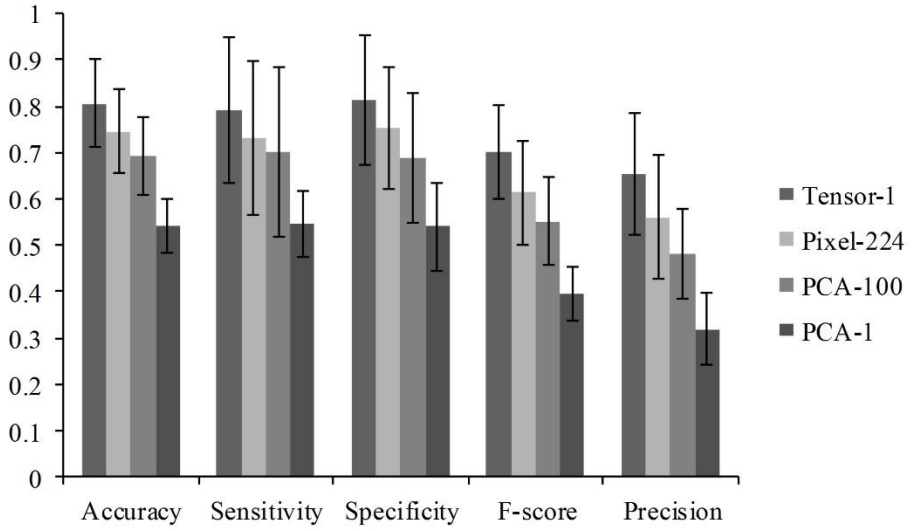


Figure 14: The performance of the Tensor-1, PAC-1, PAC-100, Pixel-224 classification methods. See the text for the definitions of the four methods.

3.4.5 Classification Results of the Spectral-Spatial Method

The spectral-spatial classification algorithm was implemented in MATLAB (Version R2013a, Mathworks, Matick, MA) using a high-performance computer with 128 GB RAM and 32 CPU cores operating at 3.1 GHz. The region of interest (ROI) selected in each mouse image was of the dimension 435 × 390. So each testing image ROI consisted of 168,004 pixels. Since surrounding tissue areas in the selected ROI were generally larger than the tumor region, the same number of healthy tissue pixels as the number of tumor pixels was randomly chosen from surrounding tissue to build a balanced training dataset.

After the completion of the SVM training process with optimal parameters, it usually takes about 2 minutes to test the hypercube image data of one mouse. The computation performance can be improved by implementing the algorithm in parallel computing and by using C++ language.

Figure 15-18 shows the SVM probability maps of all 12 tumor hypercubes in the first row and the refined classification results by active contour in the second row. As can be seen from these figures, SVM probability maps exhibited salt and pepper appearances, while active contour reduced both the false positives and false negatives in the normal tissue regions.

Mice # 889, # 893, and # 894 in Figure 15 and Mouse # 895 in Figure 16 were well classified with an average accuracy, sensitivity, and specificity of 96.1%, 93.2%, and 97.5%, respectively. Mouse # 897 and Mouse # 885 in Figure 16 exhibited misclassification in the blood vessels within normal tissue region. The binary classification only sorts the pixels into two categories: tumor or normal tissue. Blood vessels showed strong haemoglobin signals which might appear closer to the tumor regions, and therefore were misclassified as tumor pixels.

Mice # 892, # 898, and # 888 in Figure 17 contained false positives in the normal tissue which was curved and highly resembled the tumor tissue. Mice # 896, # 890, and # 891 in Figure 18 were not classified satisfactory, which was consistent with the observations from the GFP images in Figure 12. Due to the random scattering caused by the tissue artifacts, the reflectance spectral differences between cancerous and normal tissue were obscured. Therefore, the classification algorithm was not able to achieve satisfactory results.

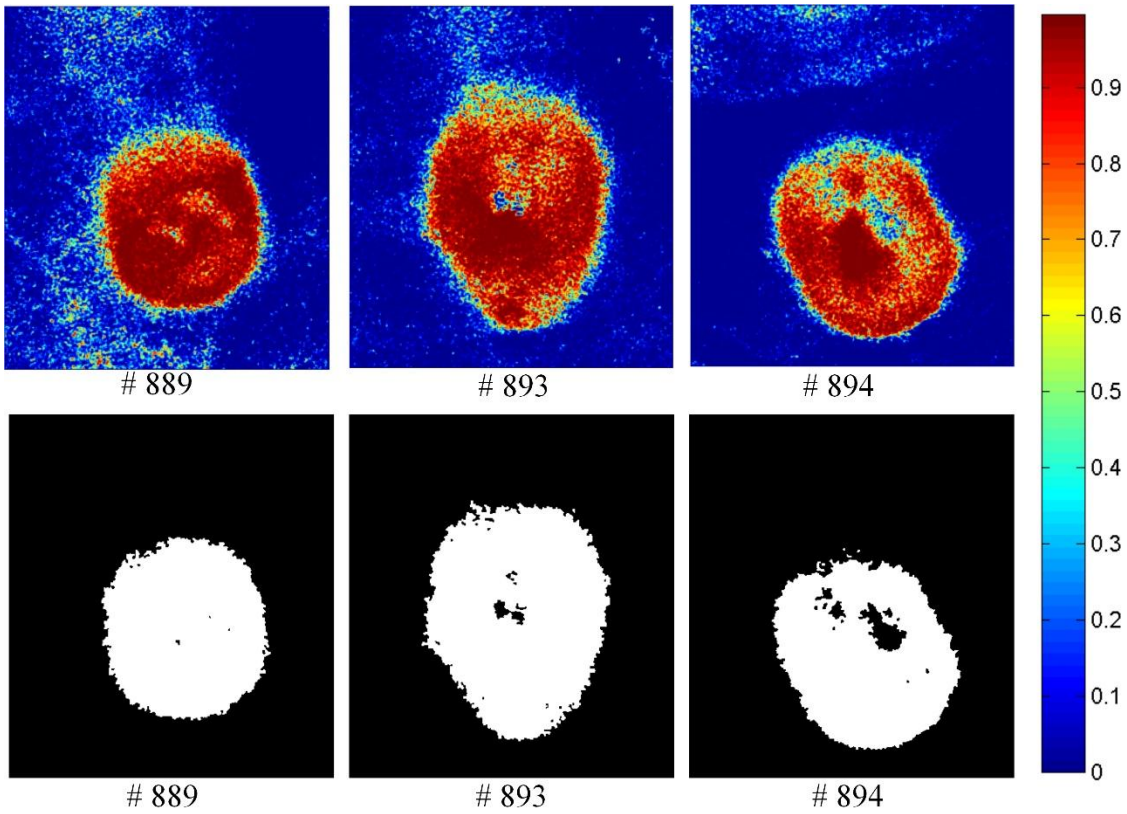


Figure 15: Classification maps. The first row represents the SVM probability map for Mice # 889, # 893, # 894, and the color bar on the right denotes the probability with different colors. The second row represents the corresponding binary tumor maps after active contour post-processing.

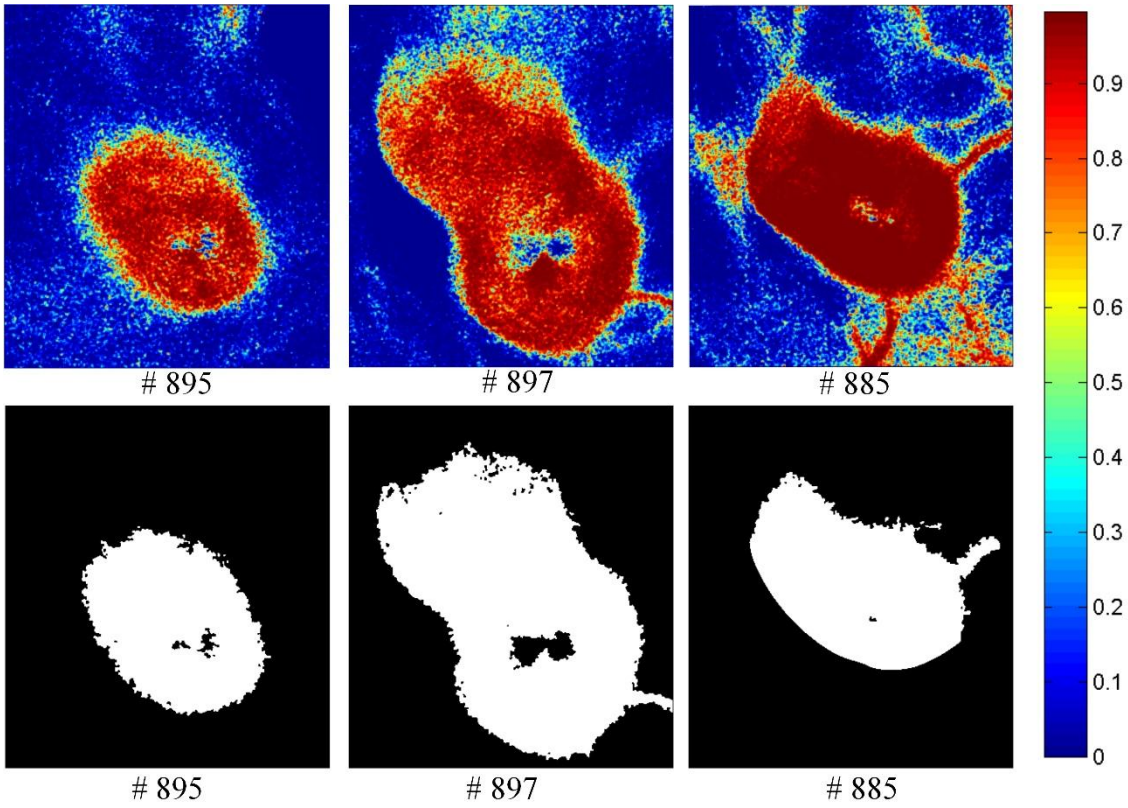


Figure 16: Classification maps. The first row represents the SVM probability map for Mice # 895, # 897, # 885, and the color bar on the right denotes the probability with different colors. The second row represents the corresponding binary tumor maps after active contour post-processing.

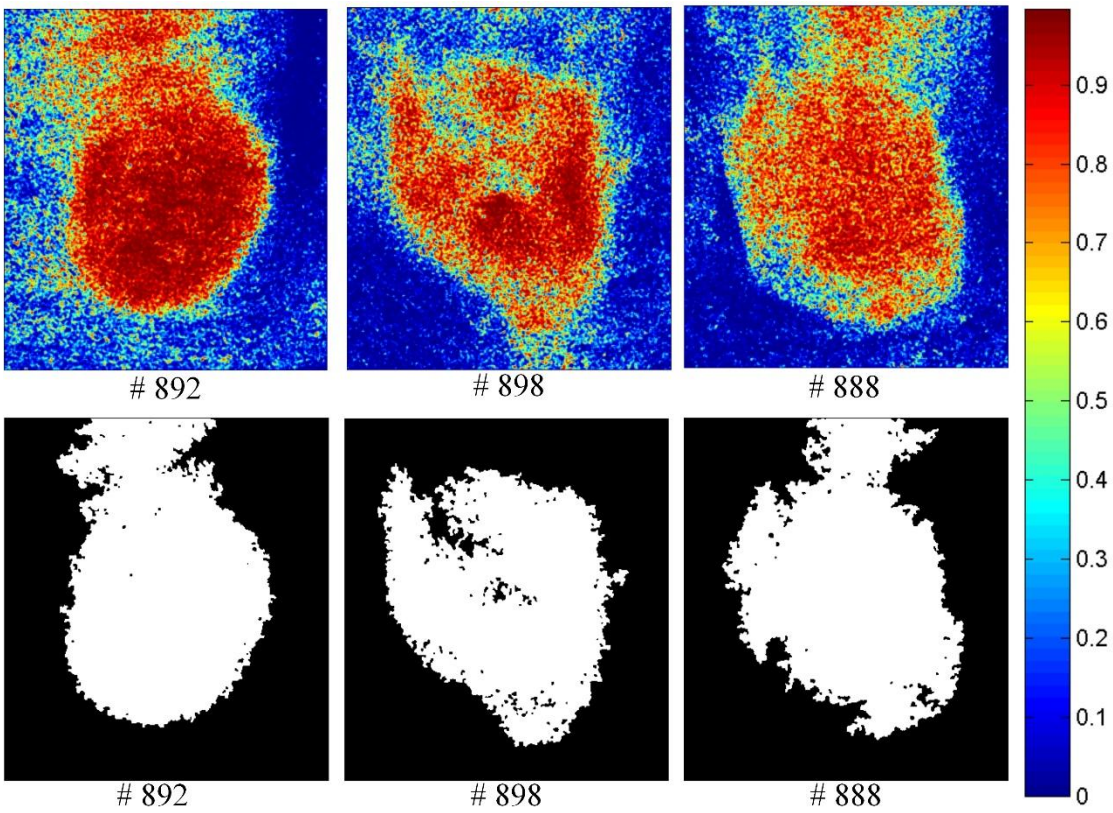


Figure 17: Classification maps. The first row represents the SVM probability map for Mice # 892, # 898, # 888, and the color bar on the right denotes the probability with different colors. The second row represents the corresponding binary tumor maps after active contour post-processing.

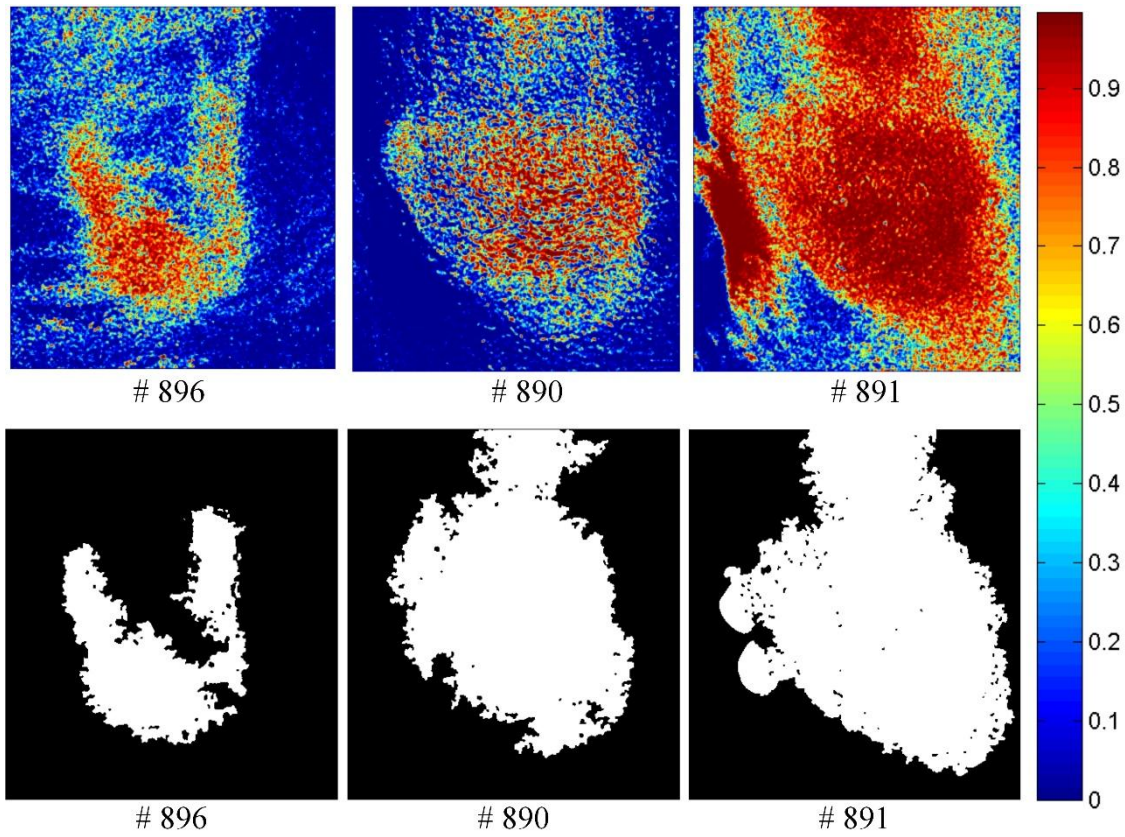


Figure 18: Classification maps. The first row represents the SVM probability map for Mice # 896, # 890, # 891, and the color bar on the right denotes the probability with different colors. The second row represents the corresponding binary tumor maps after active contour post-processing.

Figure 19 shows the comparison of the classification performance before post-processing and after the processing. The average accuracy, sensitivity, specificity, F_score, and precision of all twelve mice with standard deviations as error bars were plotted. The active contour post-processing procedure further improved the average accuracy, sensitivity, and specificity of the twelve mice by 5.1%, 1.6% and 6.7%, respectively.

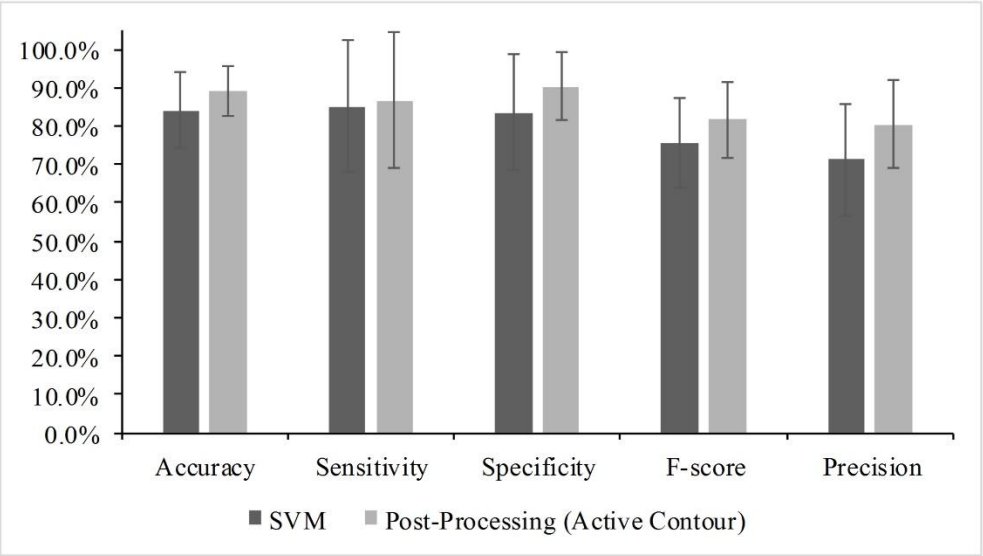


Figure 19: Comparison between the classification results before and after post-processing. As can be seen, the active contour post-processing method improved the classification performance of SVM in all the metrics.

The classification performance for all mice after post-processing is listed in Table 2. The average accuracy, sensitivity, and specificity of 12 mice were 89.1%, 86.8%, and 90.4%, respectively. Based on the above analysis, we knew that the bad performance in the last three mice shown in the shaded area of Table 2 was caused by the tissue artifacts. If we remove these three mice, the average accuracy, sensitivity, and specificity were 91.9%, 93.7%, and 91.3%, respectively.

Table 2: Summary of the classification performance

Mice ID	Accuracy	Sensitivity	Specificity
# 889	97.1%	94.9%	97.7%
# 893	96.9%	96.0%	97.3%
# 894	95.8%	88.7%	97.7%
# 895	94.5%	96.3%	94.2%
# 897	91.4%	91.3%	91.5%
# 885	90.9%	99.8%	88.6%
# 892	85.0%	99.9%	80.0%
# 898	88.3%	92.4%	86.6%
# 888	86.8%	84.2%	88.4%
# 896	85.2%	55.0%	96.2%
# 890	79.1%	46.2%	97.9%
# 891	78.4%	96.7%	68.3%
Mean	**89.1%**	**86.8%**	**90.4%**
Std	**6.5%**	**17.6%**	**8.9%**

3.5 Discussion

This paper has described and evaluated a spectral-spatial classification method for distinguishing cancerous and healthy tissue *in vivo* in a head and neck cancer animal model and demonstrated that hyperspectral imaging combined with a spectral-spatial classification method holds great promise for the noninvasive diagnosis of cancer.

The basis for cancer detection using hyperspectral imaging and the spectral-spatial method arises from the differences in the spectra obtained from the normal and diseased tissue due to the multiple physiological changes associated with tissue

transformation from healthy to cancerous stages. During disease progression, healthy tissue transforms to pathological tissue with biochemical and morphological changes, such as increase in epithelial thickness, nuclear size, nuclear to cytoplasmic ratio, changes in the chromatin texture and collagen content, and angiogenesis [175]. These changes modify the diffusely reflected light, therefore reflectance spectra exhibit spectral features associated with different biochemical and morphological characteristics of cancerous and normal tissues.

The motivation for developing spectral-spatial classification algorithms is that hyperspectral imaging is limited by user's ability to pull relevant information out of the enormous amount of data, and the development of advanced data mining methods utilizing the abundant spectral and spatial information contained in hypercube are desirable for classification of lesions and healthy tissue. As pre-clinical and clinical research with hyperspectral imaging moves forward, more and more dataset are going to be acquired and stored. On the one hand, spectral databases for different types of pathological tissues, cells, and molecules could aid in a better interpretation of hyperspectral images. On the other hand, the hidden patterns, unknown correlations, and useful diagnostic information could be uncovered and fully utilized with the help of machine learning and data mining methods. For example, classification models build upon large dataset would provide a valuable tool for quantitative diagnosis of cancer.

The prominent advantage of hyperspectral imaging is that it combines the wide field imaging with spectroscopy. Additionally, hyperspectral imaging is a noninvasive, non-ionizing imaging technology, which does not require contrast agents. Hyperspectral images have more spectral channels and higher spectral resolution than RGB images, which might carry more useful information for characterization of physiology and pathophysiology. Spectroscopy measures tissue point by point, which might miss the most malignant potential. Hyperspectral imaging captures the spectral images of a large

area of tissue, which overcomes the under-sampling problem associated with spectroscopy and biopsy.

The application of hyperspectral imaging is limited because it examines only the areas of tissue near the surface, which could be a big problem for imaging deep resided tumors inside tissue *in vivo*. The optical penetration depth is defined as the tissue thickness that reduces the light intensity to 37 percent of the intensity at the surface. Bashkatov et al. [176] measured the optical penetration depth of light into skin over the wavelength range from 400 to 2000 nm. It was observed that light penetration depth at wavelength 450 nm was about 0.5 mm, and light at above 500 nm had a penetration depth of above 1 mm. The maximum penetration depth was found to be 3.5 mm at wavelength 1090 nm. Given that the skin of the nude mice in our experiment was less than 1 mm, and that the reflectance images starting from 450 nm already showed the vascular beds of tumor, it makes sense to assume that the reflectance spectra acquired from 450 to 900 nm carries diagnostic information about tumors underneath the skin. Therefore, hyperspectral reflectance imaging combined with spectral-spatial classification methods can distinguish between tumors and surrounding tissue through intact skins.

To further extend the application of the proposed technique, we plan to explore the ability of hyperspectral imaging at the near-infrared region above 900 nm for noninvasive cancer detection in the future because near-infrared light has relatively deep tissue penetration compared to visible light. Since the near-infrared spectrum is complicated by the presence of overlapping water bands and vibrational overtones, it becomes more difficult to interpret the large volume of hyperspectral dataset by decomposing it into different tissue components. Therefore, spectral-spatial classification based on the spectral differences of cancerous and healthy tissue pixels would be of great importance for the statistical analysis of hypercubes.

The computational requirement for handling the vast amount of hyperspectral data is demanding in terms of the computer resources and time costs. Motion artifacts may

also deteriorates the image quality due to longer data acquisition times for large dataset. Therefore, optimal band selection will be performed before applying the spectral-spatial classification method in our future research. Once the reflectance spectral bands which best characterize the tissue physiology were selected, only spectral images at specific wavelengths will be acquired and used for further analysis.

3.6 Conclusion

In this study, we described and validated a spectral-spatial classification framework based on tensor modeling for hyperspectral imaging in the application of head and neck cancer detection. Third-order tensor provided a natural representation for a hypercube and Tucker tensor decomposition preserved the most discriminative information contained in each hypercube for cancer detection. Tensor-based method characterized both spatial and spectral properties of the hypercube and performed dimensionality reduction effectively. The proposed classification method was able to distinguish between tumor and normal tissue in an animal head and neck cancer model. The results from this study suggested that the combination of hyperspectral imaging with spectral-spatial classification methods may enable the detection of head and neck cancers in a noninvasive manner.

CHAPTER 4

DETECTION AND DELINEATION OF SQUAMOUS NEOPLASIA WITH HYPERSPECTRAL IMAGING IN A MOUSE MODEL OF TONGUE CARCINOGENESIS

Oral cancer has a tendency to be detected at a late stage, which is detrimental to the patients because of its high mortality and morbidity rates. Early detection of oral cancer is therefore important to reduce the burden of this devastating disease. In Chapter 3, we demonstrated the feasibility of using HSI for noninvasive cancer detection in a subcutaneous xenograft tumor model. To further evaluate the potential of HSI as a diagnostic tool for oral tumors, an animal study was designed to acquire hyperspectral images of *in vivo* and *ex vivo* mouse tongues from a chemically-induced tongue carcinogenesis model. The pathological maps were reconstructed to match the dorsal surface of the tongue as the gold standard for validation. A variety of machine learning algorithms, including linear and quadratic discriminant analysis, ensemble learning methods, and support vector machines, were evaluated for tongue neoplasia detection with HIS and validated by the reconstructed pathological gold standard maps. The diagnostic performance of HSI, autofluorescence imaging, and fluorescence imaging were compared in this study. Color-coded maps were generated to display the predicted location and distribution of neoplasia. The diagnostic performance of HSI for the distinction of neoplastic and non-neoplastic tongue tissue was shown to be comparable to autofluorescence imaging, and fluorescence imaging of proflavine. This study demonstrated the feasibility of HSI combined with machine learning techniques for the detection and delineation of squamous neoplasia in a chemically-induced carcinogenesis model.

4.1 Detection and Delineation of Squamous Neoplasia with Hyperspectral Image Classification

4.1.1 Introduction

Oral cancers are among the most common cancers globally, with an estimated 300,400 new cases and 145,400 deaths in 2012, thus creating a significant worldwide health problem [177]. Oral cancer can often be cured at a localized stage, but most people who develop oral cancer have advanced disease at the time of diagnosis. Among those with oral cavity and pharynx cancer in the United States, 65% have regional or distant spread of their disease at the time of diagnosis [5]. Late stage cancer identification can lead to poor survival with speech, swallowing, and cosmetic problems. The 5-year survival rate for localized oral cavity and pharynx cancer is 83.0%, but the survival rate drops to only 37.7% for metastasized cancer [5]. Patients who survive an initial occurrence of oral cancer are known to be at an increased risk of developing a second primary tumor, which is a leading cause of death in head and neck cancer patients [178]. Despite significant advances in cancer treatment, early detection of oral cancer and its curable precursors remains the most promising way to improve the survival and quality of life of patients [179].

The conventional procedure for oral cancer screening consists of visual inspection and palpation of the entire tissue surface at risk under incandescent light, followed by biopsy of highly suspicious tissue regions to make definitive diagnosis of malignancy [180]. Visual examination relies heavily on the experience and skills of the physician. In addition, it is often difficult even for an experienced clinician to differentiate malignant lesions from benign lesions or premalignant lesions and subsequently decide the area to biopsy. Due to the heterogeneous morphology and visual appearance of the lesions, biopsy diagnosis may not be representative of the highest pathological grade of a tumor due to the small sampling area [10]. In addition, biopsy is invasive, label-intensive,

subjective and can be inconsistent due to intra-observer and inter-observer variation [10] [12]. Approaches that can distinguish between normal and neoplastic tissue in a reliable and noninvasive manner would be very useful in the clinical setting to facilitate and improve early diagnosis of oral cancer, thus decreasing the morbidity and mortality of the disease.

More than 90% of malignant neoplasia of the oral cavity are squamous cell carcinoma, which originates in the oral mucosal lining of the oral cavity [6]. The direct access to the lesion in the oral cavity makes it possible to detect superficial tumors with optical imaging, which offers a variety of noninvasive tools utilizing intrinsic and extrinsic tissue contrast for early detection of oral cancer. The intrinsic approaches involve the development of optical devices to specifically explore the alterations of tissue absorption and scattering properties by reflectance imaging and to reveal levels of endogenous chromophores such as reduced nicotinamide adenine dinucleotide (NADH) and flavin adenine dinucleotide (FAD) (autofluorescence) accompanying malignant progression. The extrinsic methods rely on the exogenous molecular imaging agents that can be topically applied on the tissue surface to enhance detection of changes (fluorescence). One study reported the use of a point spectroscopy device combining intrinsic fluorescence, diffuse reflectance, and scattering signals to collect signals from multiple sites within the oral cavity, with 96% sensitivity and 96% specificity for the distinction of cancerous/dysplastic from normal tissue, and with 65% sensitivity and 90% specificity for discrimination of dysplastic from cancerous tissue [144]. Although the results were promising, only a very limited number of measurements were made from selected sites of oral cavity. In addition, it would be difficult to screen the entire oral cavity with small optic fibers, which limited the clinical use of spectroscopy for oral cancer screening. In another study, multispectral imaging with reflectance wavelengths correspond to hemoglobin spectral features was shown to maximize microvasculature visualization and contrast of oral tissue [181]. However, only two wavelengths, 530 nm

and 600 nm, were explored, without quantitative evaluations of their diagnostic performance.

Two fluorescence vital-dyes have been previously tested for the detection of oral neoplasia [182]. One of the vital dyes is 2-deoxy-2-[(7-nitro-2,1,3-benzoxadiazol-4-yl)amino]-D-glucose (2-NBDG, Cayman Chemical, Ann Arbor, Michigan), which is a fluorescent deoxyglucose molecule that can be used to assess the metabolic activity of cells. The other fluorescence dye is proflavine, which can non-specifically stain cellular structures and allows for visualization of nuclear morphology. The mean fluorescence intensity of 2-NBDG and the standard deviation of the proflavine intensity have been reported to discriminate neoplasia (moderate dysplasia, severe dysplasia, and cancer) from non-neoplastic regions of interest with 91% sensitivity and specificity [182]. However, this method was shown to be only 42% sensitive when applied to classify mild dysplasia.

We have investigated the feasibility of using HSI for cancer detection in a subcutaneous mouse model of prostate cancer and that of head and neck cancers (HNC). With a spectral-spatial classification method for hyperspectral images from 450 nm to 950 nm, we were able to achieve an average sensitivity and specificity of 87% and 90% for the noninvasive detection of HNC. However, there are two major limitations with our previous animal studies. The first limitation is that subcutaneous models cannot reproduce all stages of cancer. In addition, green fluorescence protein, which doesn't exist in human cells, was utilized to delineate *in vivo* gold standard for validation of detection results. The second limitation is that tumor volumes in this animal model were relatively large and clearly visible by the eyes (95 mm^2~200 mm^2), which made it easier for imaging and detection.

To overcome the above limitations and further evaluate the capability of HSI for HNC detection, we designed a longitudinal study for intrinsic and extrinsic HSI of a chemically-induced oral carcinogenesis in mice. 4-Nitroquinoline-1-oxide (4NQO),

4NQO, a synthetic water-soluble carcinogen, has been widely used in murine models to study all stages of oral carcinogenesis [183]. One advantage of this model is that 4NQO-induced lesions exhibit similar histological and molecular changes as observed clinically in human oral carcinogenesis [183]. This animal model mimics human oral neoplastic transformation with reproducible isolation of all stages including dysplasia, carcinoma in situ (CIS), and squamous cell carcinoma (SCC), therefore providing an excellent opportunity to investigate the application of HSI for the noninvasive detection of squamous neoplasia.

The goal of our study is to develop supervised learning methods for the distinction of neoplasia (dysplasia, CIS and SCC) and non-neoplastic tongue tissue, and validate the diagnostic performance of HSI by histopathology gold standard. Since HSI can provide images relying on intrinsic tissue contrast alone (reflectance, autofluorescence) or using exogenous contrast agents mapping the expression of biomarkers (fluorescence), we compare the diagnostic performance of label-free HSI with autofluorescence and vital-dye fluorescence imaging of 2-NBDG and proflavine for the detection and delineation of squamous neoplasia.

4.1.2 Instrumentation

The hyperspectral imaging system used in this study was described in Chapter 2.

4.1.3 Mouse Tongue Carcinogenesis Model

Six-week-old female CBA/J mice (Jackson Laboratoray, Bar Harbor, ME) were housed in the animal care facility under controlled conditions. 4NQO powder (Sigma Aldrich, Saint Louis, USA) was dissolved in the drinking water of the mice. The solution was stored at 4 ºC and a fresh aliquot was applied and changed weekly. In this experiment, thirty mice were divided into an experimental group (N = 24) and a control group (N = 6). In the experimental group, mice were treated with 4NQO solution (100

μg/ml) in their drinking water for 16 consecutive weeks to induce epithelial carcinogenesis. In the control group, mice received normal drinking water without 4NQO. All the animals were fed with sterilized special diet (Teklad global 18% protein rodent diet, Harlan, Indianapolis, Indiana) and autoclaved water throughout the study period. The mice in each group were evaluated weekly for body weight and water consumption. The experiment was terminated at 24 weeks. All of the animal procedures were conducted in accordance with the Guidelines for the Care and Use of Laboratory Animals, and were approved by Institutional Animal Care and Use Committee (IACUC) of Emory University.

4.1.4 Data Acquisition

Mice were scanned with hyperspectral reflectance imaging once a month for up to 24 weeks or until signs of sickness or weight loss. At week 8, 12, 20, and 24 mice were randomly selected from both groups for hyperspectral imaging. The entire experimental procedures were summarized in Figure 20 and described below:

(1) Before imaging the animals, we acquired white and dark reference hypercubes. The white reference hypercube was acquired by imaging a standard white reference board in the field of view. The dark reference hypercube was acquired by keeping the camera shutter closed.

(2) After the reference hypercubes were acquired, we anesthetized the selected mouse with ketamine. In order to image the *in vivo* tongue, we placed the mouse in a supine position, gently pulled out the mouse tongue, and taped it on a sterilized imaging stage.

(3) During *in vivo* imaging of the tongue, we first acquired reflectance hyperspectral images from 450-950 with 5nm intervals. Next we acquired autofluorescence images of the tongue with 455nm excitation and 490 nm long-pass emission filter.

(4) After *in vivo* imaging was complete, we euthanized the mouse by cervical dislocation and collected the tongue tissue for immediate *ex vivo* imaging, including reflectance

hyperspectral imaging, autofluorescence imaging, and fluorescence imaging. Reflectance and autofluorescence imaging procedures were the same as in step (3). Fluorescence imaging methods for 2NBDG and proflavine were previously described in [182]. Here is a brief summary of the imaging procedures:

2-NBDG Imaging: The tissue specimen was incubated in a 160 μM solution of 2-NBDG (Cayman Chemical, Ann Arbor, Michigan) in 1× phosphate-buffered saline (PBS) for 20 min at 37°C. Next the specimen was washed once with PBS and fluorescence images of 2-NBDG were acquired using the blue excitation and 490 long-pass emission.

Proflavine Imaging: The specimen was incubated in a 0.01% w/v solution of proflavine (Sigma Aldrich, St. Louis, Missouri) in 1× PBS for 2 min at room temperature. Next the specimen was washed once with PBS and fluorescence images of proflavine were acquired using blue excitation and 490 nm emission. The fluorescence signal from proflavine staining is much brighter than that of 2-NBDG, allowing for imaging of proflavine-stained tissue after 2-NBDG staining.

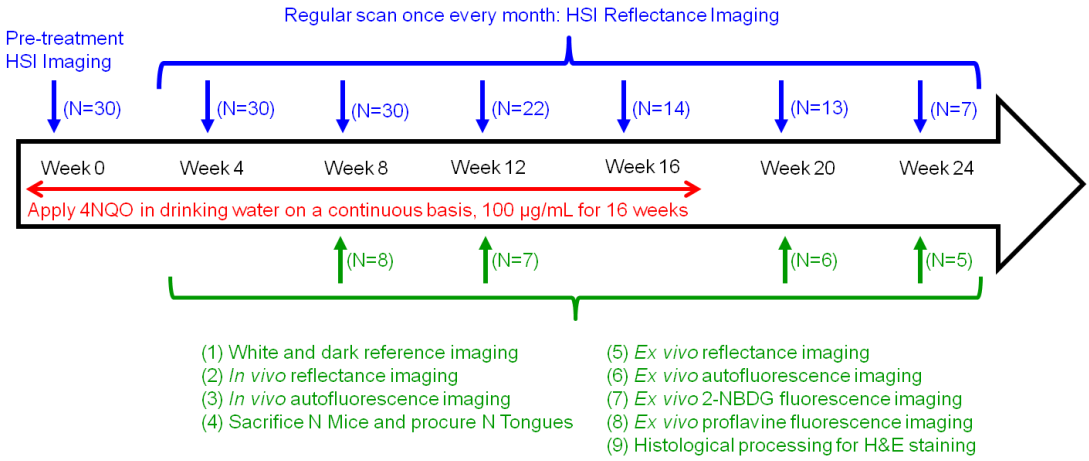


Figure 20: Hyperspectral imaging of 4NQO-induced mouse tongue carcinogenesis model.

4.1.5 Histology Correlation

Immediately after *ex vivo* imaging of the dissected tongue, the ventral surface was inked with blue on the left side and red on the right side, which would aid in the identification and correct orientation of the tongue during subsequent processing. The inked tongue specimens were placed in 10% buffered formalin overnight for fixation. Each fixed tongue was further processed through dehydration, clearing, and wax infiltration. Next, the specimen was embedded in a cassette to form a paraffin block, which was then clamped into a microtome for tissue sectioning. Tongue tissues procured at week 8 and 12 were sectioned sagitally into a series of 5 μm slices. Specimens procured at week 20 and 24 were first bisected longitudinally along the midline groove and then embedded in paraffin blocks, which were sagitally sectioned into a series of 5 μm slices. The interval between two tissue sections was 200 μm for week 12, and 100 μm for week 20 and 24. Histological slides were stained with H&E and then digitized for pathological diagnosis.

Pathology Diagnosis

The H&E slides were reviewed by an experienced pathologist specializing in head and neck cancer (Susan Muller), who graded each tongue slice by the most severe pathology it contains. Next, the dorsal surface of each tongue slice was further segmented into pathological regions of normal (including inflammation and hyperplasia), dysplasia, CIS, and SCC as demonstrated in Figure 21. Four tissue slices were laid onto each glass slide immediately after sectioning. Some tissue slices close to the end of the glass slide may not show up on the digitized slides due to the limited scanning region, which were labeled as missing gold standard. The grading and diagnosis of tongue cancer is based on the presence of architectural and cytological changes of epithelial layer based on microscopic examination of H&E stained sections [184].

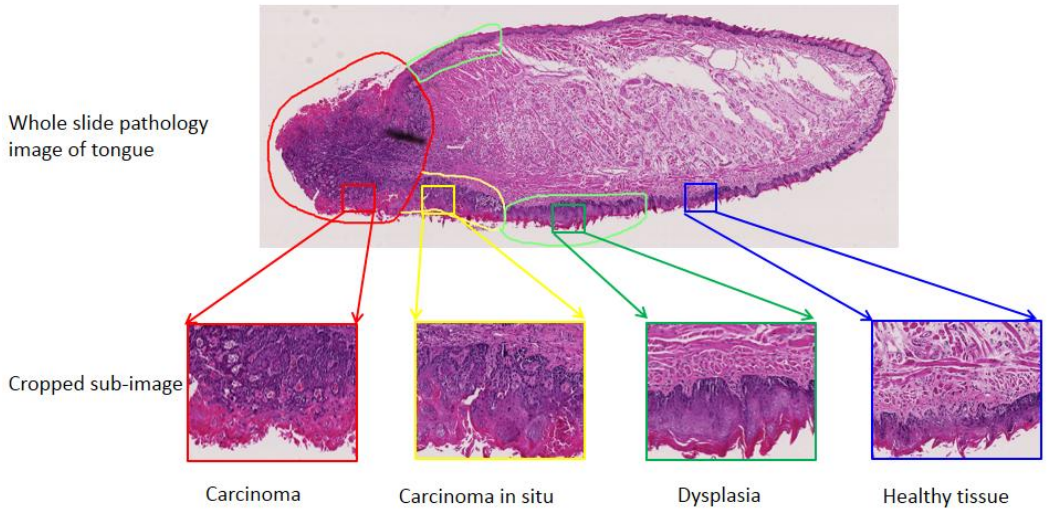


Figure 21: An example of pathology grading and annotation in a whole slice image of tongue specimen.

Pathology Gold Standard Map Generation

The first challenging task in this study was to generate pathology gold standard maps that match with the dorsal surface of the tongue to validate the cancer detection results from hyperspectral images. This was not a simple registration problem between hyperspectral images and pathological images; because the former are the projections of dorsal tongue surface while the latter are the cross-sections of the tongue with both dorsal and ventral surfaces. To reconstruct the pathology map, we tracked each step of the histological processing to ensure correct tongue orientation during embedding, numbering of each tissue slice, and specific layout of tissue slices on each glass slide during sectioning. As shown in Figure 22, each tongue slice was sectioned sagitally into a series of 5 µm tissue slices with intervals of 100 µm or 200 µm between sections. Each H&E slice represented one straight line parallel to the midline on the dorsal surface of the tongue. The length of the midline of a tongue slice and that of different pathological regions within the dorsal surface of the slice were measured. The pathology readings for

each tongue can then be represented as a matrix with each row corresponding to one tongue slice and each column representing the length ratio of individual pathology regions to the length of the tongue. In this way, the pathology readings on the dorsal surface of the tongue slice could be mapped back proportionally onto the corresponding region of the tongue HSI image. A 2D color-coded pathology map was created and aligned with the hyperspectral images of each tongue as shown in the last image of Figure 22, where the red line indicates SCC region, the yellow line indicates CIS, green dysplasia, and blue healthy region. Some slices were labelled as black lines on the pathology map indicating missing gold standard.

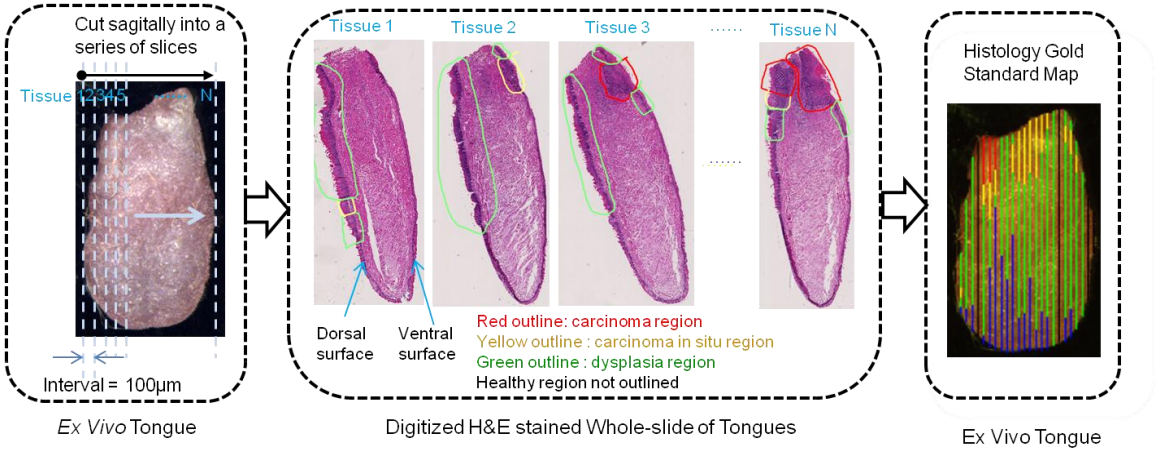


Figure 22: Schematic diagram of pathological gold standard map generation

4.1.6 Hyperspectral Image Classification

Hyperspectral Image Pre-processing

The pre-processing of hyperspectral imaging consists of two steps. The first step was to convert the raw data into percent reflectance value in a pixel-wise manner using Equation (1). The second step was to detect and remove glare pixels from the normalized

hyperspectral images. Glare regions were formed due to specular reflection from the moist tongue surfaces, which contained no diagnostic information under the tissue surface. Glare pixels can introduce artifacts in feature extraction and deteriorate classification results. We observed that glare pixels were characterized with very bright reflectance intensity; therefore they generally fell into the long tail region in the intensity histogram of the sum image for all the spectral bands. To identify a threshold to detect these pixels, we developed an adaptive thresholding method, which fitted the histogram of the sum image with a loglogistic distribution and took a certain percentage (α) of the peak intensity in the loglogistic distribution. The loglogistic distribution fitted the histogram very well. The value of α was experimentally set to 0.05% through trial and error, which was found sufficient to detect all glare pixels.

Feature Extraction

Reflectance Hyperspectral Imaging

For reflectance hyperspectral image data, we evaluated and compared pixel-wise and block-based feature extraction (FE) methods: (1) pixel-wise spectral method extracts the normalized reflectance intensity of each pixel on a straight line of the pathology map and each pixel was labeled a pathology type. (2) block-based spectral method first grids each tongue hypercube ("M×N×K" , "M×N" is the size of each spectral image and "K" is the number of wavelengths in each hypercube) into many blocks of size "m×n×K" and then averages the spectral curves of individual pixels within each block as the spectral features of this block. Each block is labeled with the most frequent pathology type in this block. In order to examine the contribution of different wavelength regions, we also extracted spectral features from different sub-regions of 450-950nm, including the visible (VIS) wavelength region of spectral curve ranging from 450-600 nm, 605-850 nm, and the near-infrared (NIR) wavelength region of 855-950 nm.

Autofluorescence Imaging

For autofluorescence imaging, we extracted the average fluorescence intensity from 500 to 720 nm with 5 nm increments within each block for classification. It has been reported that red-to-green fluorescence at 405 nm excitation is effective in discrimination between neoplastic and non-neoplastic areas of the oral cavity [185]. So we also extracted the red-to-green (red: 650 nm, green: 510 nm) fluorescence for comparison.

Vital-dye Fluorescence Imaging

For 2NBDG and proflavine imaging, we extracted the mean fluorescence intensity of 2-NBDG or proflaving staining from 500 to 720 nm with 5 nm increment within each block for classification. It has been shown that the mean of 2NBDG and the standard deviation of proflavine could separate neoplastic from non-neoplastic regions of interest with 91% sensitivity and specificity [182]. The emission maximum of 2NBDG is around 540 nm with 450 nm excitation, so the mean fluorescence intensity at this band was extracted for comparison. In addition, the peak emission is around 515 nm with 450 nm excitation, so the standard deviation of fluorescence intensity at 515 nm is used as another feature.

Supervised Classification

The goal for supervised learning is to distinguish neoplastic from non-neoplastic tissue from both *ex vivo* and *in vivo* tongue. The neoplastic class included lesions histopathologically diagnosed as dysplasia, CIS, and SCC, while the non-neoplastic class included clinically normal sites of the tongue.

To build a prediction model with low bias and low variances, we tested a variety of classifiers including linear discriminant analysis (LDA) [64], quadratic discriminant analysis (QDA), ensemble LDA [70], RUSBoost [71], random forests (RFs) [72], linear

support vector machines (SVMs), and kernel SVMs with radial basis function (RBF) [88]. The basic principles and characteristics of these classifiers were summarized in Chapter 2:

Performance Evaluation

We assess the performance of classifiers with a variety of performance metrics including receiver operating characteristic (ROC) curves, the areas under the ROC curve (AUC), accuracy, sensitivity, and specificity. The ROC curve offers a graphical interpretation of the trade-off between sensitivity and specificity for a range of possible cut-off points. AUC of a classifier is equivalent to the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance. At the optimal operating point of ROC curve, we can obtain the accuracy, sensitivity, and specificity from the confusion matrix as defined in [149] [186].

4.1.7 Results

Mouse Tongue Carcinogenesis Model

Over the course of neoplastic progression, a number of mice from the 4NQO-treated group and the control group were selected for hyperspectral imaging and killed after imaging at various periods of observation. Mice in the control group were healthy throughout the experiment, while mice exposed to 4NQO with observations up to 24 weeks developed a spectrum of tongue lesions. The diagnosis was determined by the worst histologic lesion identified in the tongues of individual mice treated. At 8 weeks following the application of 4NQO, 6 out of 6 mice developed epithelial dysplasia of the tongue. At 12 weeks, CIS was detected in 2 out of 6 mice at 12 weeks and in 4 out of 6 at 20 weeks. SCC was detected in 2 out of 6 at 20 weeks, and 3 out of 3 mice at 24 weeks. Exposure to 4NQO for up to 16 weeks caused eterogeneous lesions in the tongue of mice as well as esophageal.

Predictive Analysis

Ten mice tongues meeting the following criteria were selected for quantitative analysis: (1) Have detailed gold standard of the epithelium for the dorsal tongue outlined by the pathologist. (2) Have all six types of images available, including *in vivo* reflectance hyperspectral images and autofluorescence images, and *ex vivo* reflectance hyperspectral images, autofluorescence, 2NBDG, and proflavine fluorescence images. (3) Good cutting without severe damage to the tongue regions. (4) Good imaging quality without significant motion artifacts. All these mice developed heterogeneous lesions on different sites of tongues following exposure to 4NQO for up 16 weeks. To select the optimal parameters and evaluate the performance of predictive models, we conducted nested cross validation (CV) with k-fold internal CV on the training dataset and leave-one-out external CV to estimate the model accuracy. For each fold of the external CV, we built a training model on the data from nine out of ten mice and tested the model on the remaining mice data. On each of the training dataset, we perform an additional k-fold CV to search for optimal model parameters. For ensemble LDA, the optimal number of weak learners used in the ensemble is determined by five-fold CV on a training dataset. For RUSBoost, we conduct five-fold CV to choose the minimal leaf size per leaf node and the maximal number of branch node splits per tree, as well as the optimal number of trees for the ensemble model. For linear SVMs, we conduct a 3-fold CV to search for the optimal cost parameter $C \in 2^{[-10,-9,\ldots 0,\ldots 9,10]}$. For kernel SVMs, we use a Gaussian RBF kernel with a 3-fold CV and grid search to find the optimal combination of cost parameter C and kernel parameter γ over the range of $C \in 2^{[-5,-4,\ldots 0,\ldots 4,5]}$ and $\gamma \in 2^{[-5,-4,\ldots 0,\ldots 4,5]}$. First, we compared the predictive power of pixel-wise feature extraction and block-based feature extraction with all spectral bands used for *ex vivo* tongues. The size of a block was defined as 5×5, therefore the number of samples for training and testing in block-based method was about four times less than the number of

samples in the pixel-based method, thus reducing computation time significantly. In addition, the block-based method was more accurate than the pixel-wise method in terms of all the performance metrics and classifiers. This is likely due to the fact that block-based average spectral features reduced the effect of noise and became more robust to the registration error from gold standard. Therefore, the block-based method was used in the rest of the analysis.

Next, we compared the predictive performance of seven different classification models for cancer detection in *ex vivo* tongues as shown in Table 3. Linear SVM, ensemble LDA, and LDA performed the best with an AUC value of 0.86, while SVM with RBF kernel only achieved an AUC of 0.801. LDA performed better than QDA and random forest, but slightly worse than ensemble LDA. It has been reported in [187] that LDA is more accurate than decision trees for multispectral image classification of oral cancer. RUSBoost had the highest sensitivity, but the variance was very high. In the remaining analysis, we only reported classification results with linear SVM as the classifier.

Table 3: Predictive performance of different classification models for *ex vivo* tongue cancer detection

Classifier	AUC	Accuracy	Sensitivity	Specificity
Linear SVM	0.86±0.06	79%±6%	79%±7%	79%±5%
Ensemble LDA	0.86±0.06	79%±6%	78%±7%	79%±5%
LDA	0.86±0.06	78%±6%	78%±8%	80%±4%
Random Forest	0.84±0.08	77%±7%	77%±7%	77%±8%
QDA	0.82±0.05	76%±5%	76%±6%	75%±4%
RBF SVM	0.80±0.07	75%±7%	77%±9%	75%±8%
RUSBoost	0.79±0.12	71%±15%	80%±6%	69%±20%

Wavelength Analysis of Hyperspectral Imaging

Each hypercube consists of a series of two-dimensional grayscale spectral images at wavelengths ranging from 450 nm to 950 nm. Figure 23 shows some example spectral bands from a normalized hypercube of an *ex vivo* tongue. It can be very difficult to visually identify lesions on the tongue surfaces from individual grayscale images. The image intensities vary across the wavelength region, reflecting the variations of tissue absorption and scattering properties at different wavelengths.

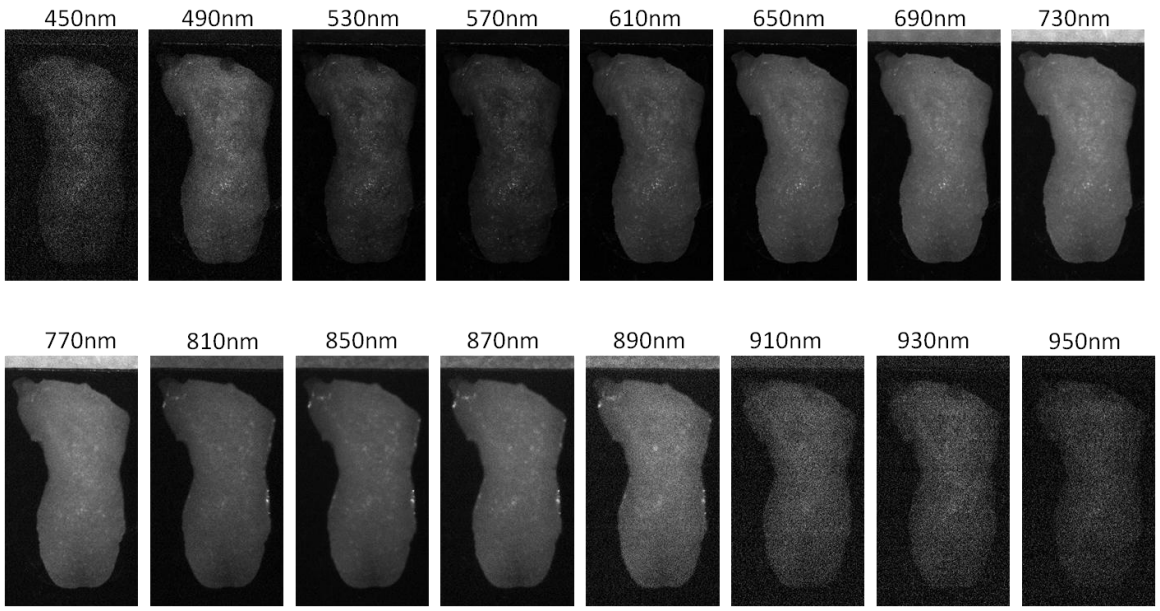


Figure 23: Example spectral bands of a hypercube for an *ex vivo* tongue.

Figure 24 shows an example of the average spectral curve of neoplastic and non-neoplastic tissue. It appears that the most prominent difference between different pathologies is in reflectance intensity. But there are many intra- and inter- variances in the spectral curves, which complicated the detection of neoplasia. Near-infrared light penetrates deeper into the tissue than visible light, thus sampling deeper tissue volumes and carrying diagnostic information from deeper tissues than visible light. Therefore, the

reflectance images in VIS and NIR light region captured by hyperspectral camera may carry complementary diagnostic information.

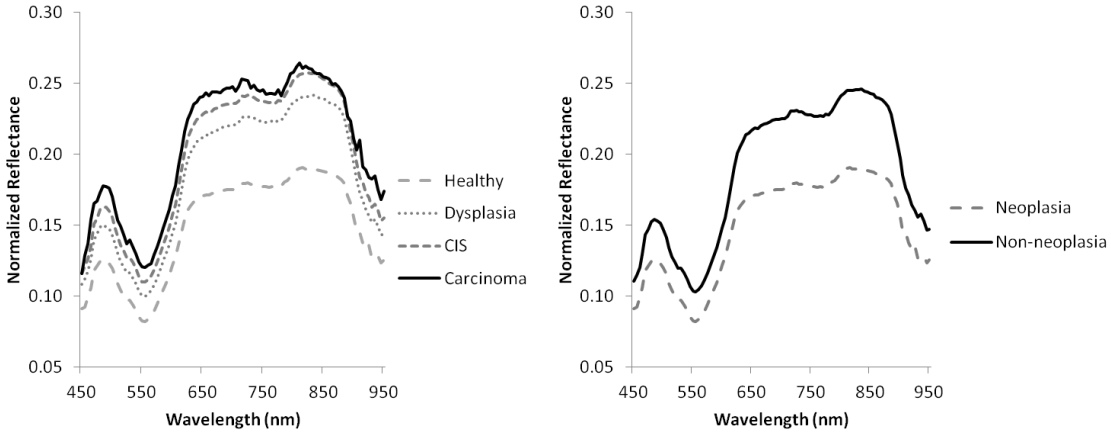


Figure 24: Example spectral curves of different tissue pathologies

To better examine the discriminatory power of different wavelength regions, we compared the diagnostic performance of 450-600 nm, 605-850 nm, 855-950 nm and 450-950 nm. As can be seen in Table 4, the whole spectrum from 450-950 outperformed the visible wavelength region of 450-600 nm and 605-850 nm, and the NIR wavelength region of 855-950 nm in both the *in vivo* and *ex vivo* tongues. This suggested that VIS and NIR light carries complementary diagnostic information for tongue cancer detection in the mouse carcinogenesis model. Since the epithelium thickness of a typical mouse tongue is only around 0.3 mm and the thickness of the tongue is around 2 mm, NIR light may have penetrated through the entire tongue. On the other hand, VIS light may only penetrate through the epithelium.The dominant chromophores in VIS region are oxygenated hemoglobin (peak absorption at 540 nm and 575 nm) and deoxygenated hemoglobin (peak absorption at 555 nm), which have been shown to be aid in visualizing increased vasculature in the oral cavity during malignant progression [188] [187]. It

should be noted that diseased human tongue can be more than 1 mm thick in the epithelium layer, with cancer cells invading into the muscle tissue, therefore NIR light could sample the cancerous areas where VIS light cannot reach and provide distinct diagnostic information for detecting neoplastic changes in the deeper tissue.

Table 4: Diagnostic performances of different wavelength regions of hyperspectral reflectance imaging for the distinction of non-neoplastic and neoplastic tissue in *ex vivo* and *in vivo* mouse tongue.

	Wavelength Region	AUC	Accuracy	Sensitivity	Specificity
In Vivo	450-600nm	0.81±0.07	76%±5%	75%±6%	75%±5%
	605-850nm	0.84±0.08	77%±6%	75%±9%	79%±6%
	855-950nm	0.83±0.08	76%±6%	74%±8%	79%±6%
	450-950nm	**0.84±0.05**	**78%±5%**	**78%±5%**	**78%±5%**
Ex Vivo	450-600nm	0.84±0.05	77%±5%	77%±6%	78%±6%
	605-850nm	0.84±0.05	77%±5%	77%±5%	78%±6%
	855-950nm	0.83±0.07	76%±7%	77%±8%	77%±8%
	450-950nm	**0.86±0.06**	**79%±6%**	**79%±7%**	**79%±5%**

Diagnostic Performances of Reflectance, Autofluorescence and Fluorescence Imaging

Hyperspectral imaging can provide images relying on intrinsic tissue contrast alone (reflectance, autofluorescence) or using exogenous contrast agents mapping the expression of biomarkers (Fluorescence). In this study, we compared the diagnostic performance of HSI (450-950nm) and autofluorescence imaging (450 nm excitation, and 490 nm longpass emission) of both *in vivo* and *ex vivo* tongues, as well as 2-NBDG fluorescence and proflavine fluorescence imaging of the *ex vivo* tongues. As shown in Table 5, among all the imaging methods, 2-NBDG fluorescence imaging of *ex vivo* tongues gave the best performance, with an average AUC, sensitivity, and specificity of 0.91, 85%, and 84% for oral lesion detection. HSI without the application of any contrast

agent achieved an average AUC, sensitivity, and specificity of 0.86, 79%, and 79%, which was comparable to the performance of proflavine fluorescence imaging of *ex vivo* tongue. For *in vivo* imaging, HSI achieved similar performance compared to autofluorescence imaging. Features including red-to-green ratio of autofluorescence intensity, mean intensity of 2-NBDG fluorescence at 540 nm, and standard deviation of proflavine fluorescence at 515 nm did not provide better prediction performance than multiband autofluorescence or fluorescence features. In general, the prediction results of *ex vivo* imaging were better than the results of *in vivo* imaging, which might be attributed to the more accurate registration of histological gold standard in *ex vivo* tongues. It was much more difficult to accurately align the *ex vivo* gold standard with *in vivo* tongues due to the stretching and motion of the tongue during *in vivo* imaging in conjunction with the tissue deformation during histological processing. Figure 25 plotted the ROC curve of neoplasia detection with HSI on individual tongues, which showed the variation of the diagnostic performance among different mice.

Table 5: Diagnostic performances of HSI, autofluorescence imaging, 2NBDG fluorescence imaging and proflavine fluorescence imaging for the distinction of non-neoplastic and neoplastic tissue of *ex vivo* tongue and the diagnostic performance of HSI and autofluorescence imaging of *in vivo* tongue.

	Imaging Method	AUC	Accuracy	Sensitivity	Specificity
In Vivo	HSI	**0.84±0.05**	**78%±5%**	**78%±5%**	**78%±5%**
	Autofluorescence	0.84±0.06	78%±5%	78%±4%	77%±9%
	Autofluorescence(R/G)	0.43±0.10	48%±6%	48%±6%	47%±12%
Ex Vivo	2-NBDG	0.91±0.04	84%±3%	85%±3%	84%±6%
	Proflavine	0.89±0.03	82%±3%	83%±3%	81%±5%
	HSI	**0.86±0.06**	**79%±6%**	**79%±7%**	**79%±5%**
	Autofluorescence	0.87±0.03	80%±4%	80%±6%	81%±3%
	Autofluorescence(R/G)	0.49±0.19	50%±14%	50%±17%	54%±12%
	2-NBDG (Mean)	0.75±0.23	70%±17%	70%±18%	71%±19%
	Proflavine (Std)	0.49±0.15	54%±9%	56%±12%	49%±20%

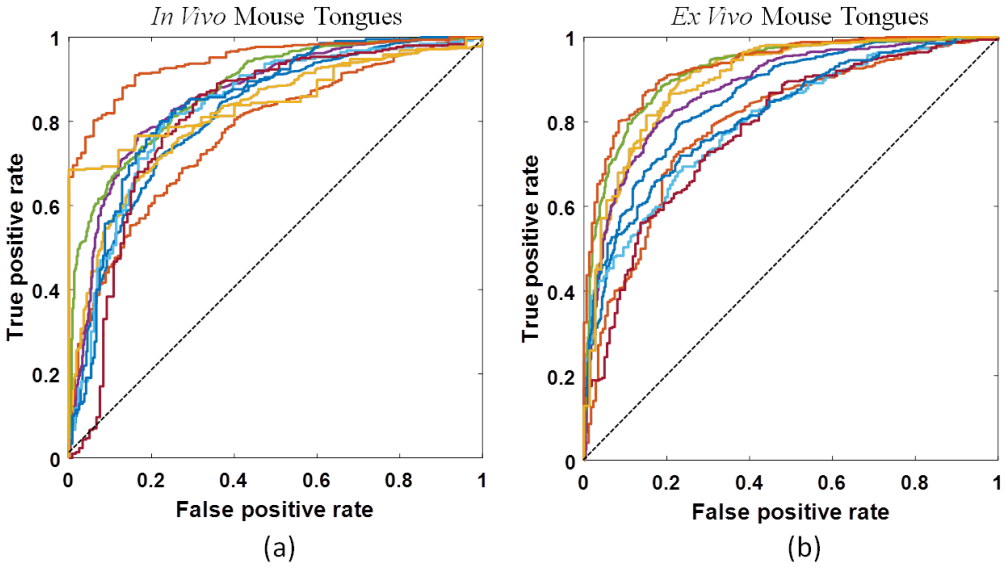


Figure 25: ROC curve of neoplasia detection with HSI on individual tongues. (a) *in vivo* mouse tongues (b) *ex vivo* mouse tongues.

In addition, we examined how many samples of dysplasia, CIS, and SCC were correctly classified. As shown in Table 6, with 2NBDG fluorescence imaging, 84% of normal tissue, 84% of dysplasia, 91% CIS, and 100.0% SCC were correctly classified. *Ex vivo* HSI was able to classify 79% of normal tissue, 79% of dysplasia, 85% of CIS, and 71% of SCC, while *in vivo* reflectance HSI was able to classify 78 % of normal tissue, 75% of dysplasia, 89% of CIS, and 81% of SCC. In general, dysplastic lesions (including mild dysplasia and moderate dysplasia) were harder than CIS to be differentiated from normal tissue due to the histological similarity of normal and dysplasia, as well as the difficulty in making accurate histological diagnosis for mild dysplasia. Only two out of ten tongues developed small areas of SCC, which made only a very small part of the dataset. Thus it was more difficult to make a robust and accurate prediction for SCC.

Table 6: Classification accuracy of different imaging method for the distinction of different pathologies

	Imaging Method	Normal	Dysplasia	CIS	SCC
In Vivo	HSI	**78%**	**75%**	**89%**	**81%**
	Autofluorescence	75%	76%	83%	91%
Ex Vivo	2NBDG	84%	84%	91%	100%
	HSI	**79%**	**79%**	**85%**	**71%**
	Proflavine	81%	82%	87%	98%
	Autofluorescence	81%	78%	88%	83%

Figure 26 showed a representative example of tongue cancer detection with *in vivo* and *ex vivo* imaging. These images illustrated the heterogeneous pathologies across the dorsal surface of the tongue, thus wide-field imaging can better capture the neoplastic changes over the entire tongue than point-based spectroscopy methods. The following sections described more details:

Hyperspectral Reflectance Imaging

It is very difficult to directly visualize the abnormal tissue transformations from the hyperspectral images of the tongue. With the proposed machine learning method, we can accurately detect and delineate the neoplastic tissue regions in the example of Figure 26 with an AUC of 0.88 and 0.92 for the *in vivo* and *ex vivo* tongues respectively. The prediction color map matched well with the gold standard color map. Both premalignant and malignant regions were accurately delineated. Most of the misclassification errors occurred in the interface of different pathologies, such as the normal and dysplasia interface, and the dysplasia and CIS interface. As seen in the gold standard map, there were finger-like protruding dysplasia regions in the interface of dysplasia and normal tongue regions, where wide-field imaging may not be as sensitive as spectroscopy. In these local regions, spectroscopy-based method could aim in the identification of the small tumor foci around the tumor-normal interface.

Autofluorescence Imaging

In human oral tissue, decreased green fluorescence when excited with UV or near-UV light has been observed, which is attributed to decreased signal from collagen crosslinks in the stroma. Autofluorescence imaging with 450 nm excitation has been shown to classify neoplastic lesions accurately [185]. Autofluorescence imaging yielded a higher AUC value for *ex vivo* tongue than for *in vivo* tongue, which might be attributed to the difficulty in registration with gold standard for *in vivo* tongue.

Fluorescence Imaging

2NBDG imaging revealed increased metabolic activity of neoplastic cells by increased fluorescence intensities, and proflavine imaging revealed the neoplastic transformation by increased and disorganized fluorescence intensity patterns from nuclei accompanying carcinogenesis. The limitation of the topically applied dye on *ex vivo* tissue was that the cut edge of the tissue or any damaged areas on tissue surface could lead to high fluorescence intensity due to nonspecific uptake, which may confound the detection of neoplastic regions. In the example case, tissue regions close to the cut edge exhibited very high fluorescence intensity, which didn't cause confusion because the cut edge region was also neoplastic.

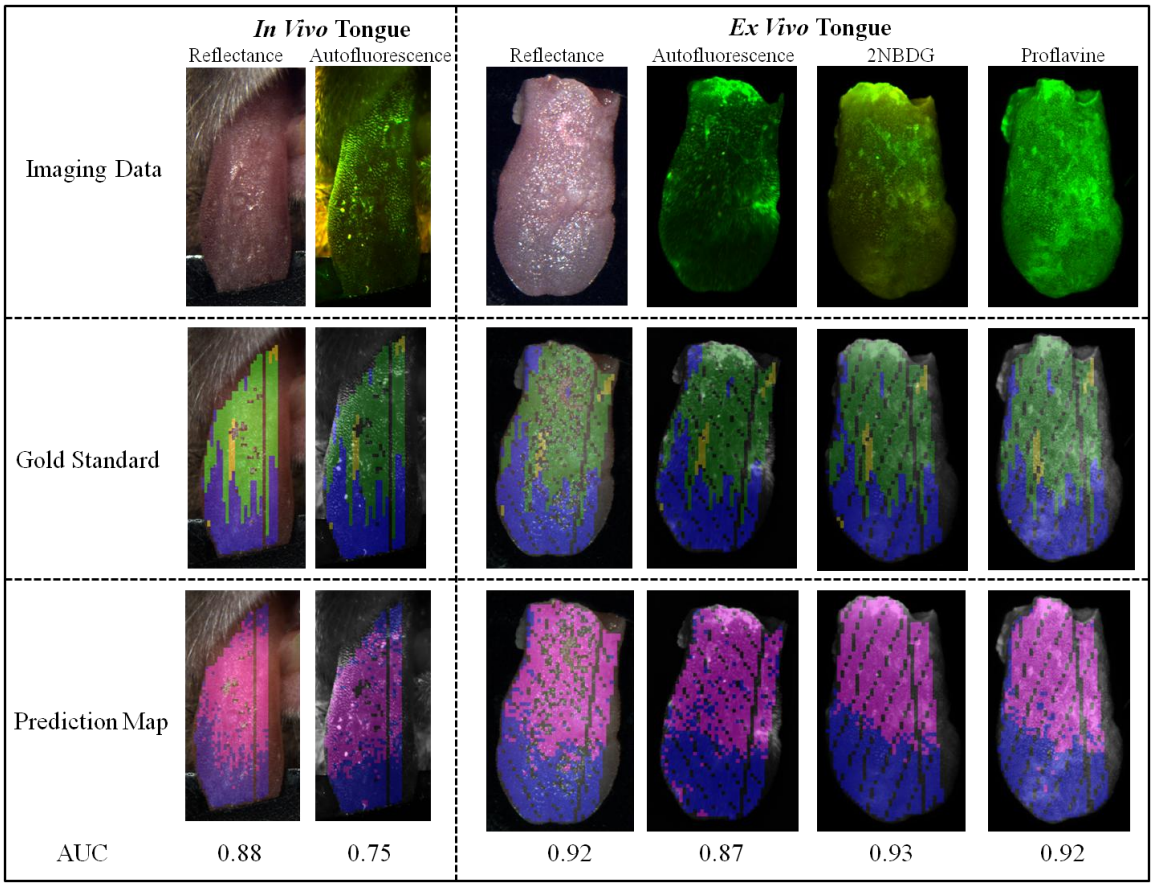


Figure 26: Example cancer prediction results for *in vivo* and *ex vivo* tongues with different imaging methods. The first row is the RGB composite images of all imaging data. The second and third rows are gold standard color maps and prediction color maps corresponding to the block-based feature extraction method respectively. In the gold standard color map, blue, green, and yellow colors indicate normal, dysplasia, and CIS respectively. In the prediction color map, magenta and cyan represent neoplastic and non-neoplastic tissue respectively. The numbers under the prediction map are AUC values of corresponding prediction models. In the gold standard and prediction map, glare pixels excluded from the feature extraction and classification as well as tissue regions missing the gold standard are left unlabeled.

4.1.8 Discussion

Screening individuals at risk for oral cancer and its precursors has the potential to improve early detection, providing the opportunity to intervene when treatment is most effective. In addition, surveillance of patients who have survived their initial oral cancer is important to identify local recurrences and second primary oral tumors, which occur at

a higher rate than for any other tumor. HSI has the potential to improve oral cancer diagnosis, decrease the use of invasive biopsy, and reduce time and discomfort associated with traditional procedures. In this study, we evaluated the diagnostic performance of HSI for the detection and delineation of tongue cancer in a chemically-induced carcinogenesis murine model. The most challenging task was to accurately register the histological gold standard with the dorsal surface of the tongue. We carefully sectioned the tongue specimen sagitally into a series of 5 μm-thick tissue slices with 100 μm or 200 μm intervals between slices. We then mapped the gold standard back onto the tongue surface proportionally along each sampling line to generate gold standard for reflectance, autofluorescence, and fluorescence images. We found the block-based feature extraction method was more accurate and robust than pixel-wise feature extraction method. Next we tested a variety of machine learning classifier, including linear and quadratic discriminant analysis, ensemble learning methods, and SVM with linear and RBF kernels, to separate neoplastic tissue (dysplasia, CIS, SCC) from normal tissue. Linear SVM, LDA, and ensemble LDA were the best classifiers among all seven classifiers tested. We also generated the final prediction map which displayed the detection and delineation of neoplastic regions compared to the gold standard color map.

While the study in [182] only examined the *ex vivo* tongue, we successfully designed an experimental method to acquire high-resolution *in vivo* images of mouse tongue while reducing the effect of motion artifacts. We used all the available image blocks rather than selected image patches, explored various machine learning algorithms, and reconstructed a prediction map that demonstrated the location and distribution of neoplasia across the heterogeneous tongue surface. In addition, our algorithm included mild dysplasia for quantitative analysis, which was a difficult histological type for classification.

The limitation of the current study was mainly in the animal model. By evaluating the diagnostic power of different wavelength region, we found it necessary to use the

whole spectrum with both VIS and NIR wavelengths for the mouse tongue model. In the case of human oral cancer detection, it would also be necessary to measure both VIS and NIR signal since tongue cancer cells can be several mm deep into the muscle, but not present in the epithelium.

To move forward to oral cancer screening in human patients, we have been working on evaluation of its diagnostic performance on fresh human surgical specimens of head and neck SCC and have achieved promising preliminary results. Design of a portable HSI instrument with higher spectral resolution and wider wavelength range into the NIR could potentially improve the diagnostic performance of HSI for oral cancer detection.

4.1.9 Conclusion

In this study, we designed an imaging experiment to acquire multimodal hyperspectral images of *in vivo* and *ex vivo* mouse tongues from a chemically-induced oral carcinogenesis model. We reconstructed the pathology map that matched well with the dorsal surface of the tongue. We implemented and validated a variety of machine learning classifiers to differentiate neoplastic from non-neoplastic tissue using HSI, compared the diagnostic performance of hyperspectral reflectance, autofluorescence, and fluorescence images, and generated prediction maps that displayed the location and distribution of neoplasia. In the future, we would like to apply HSI and machine learning technique to the diagnosis of human tongue tissue.

4.2 Histopathology Feature Mining and Association with Hyperspectral Imaging for the Detection of Squamous Neoplasia

In the previous sections, we have demonstrated that spectral signatures of tongue tissue measured by HSI contain useful diagnostic information by utilizing the entire

spectrum with machine learning methods for the detection and delineation of tongue neoplasia. The extraction of pathology-specific diagnostic information from these spectra is a complex process. Machine learning techniques can predict the disease status of the examined biological tissue with the spectral signature obtained by HSI. However, the underlying pathophysiology for the spectral difference between normal and neoplastic tissue is not well understood. In this section, we proposed to identify the association between histological features from digitized histological images quantifying the architectural features of neoplasia on a microscopic scale, with the spectral signature of the corresponding tissue measured by HSI on a macroscopic level, which may provide insight into the pathophysiology underlying the hyperspectral dataset.

4.2.1 Introduction

The gold standard for cancer diagnosis remains tissue biopsy with pathological assessment made by pathologists using visual examination of H&E stained sections under the microscope [184]. However, the effectiveness of cancer diagnosis is highly dependent on the attention and experience of the pathologists. This technique is considered invasive, expensive, and time-consuming. The diagnosis and grading of oral epithelial dysplasia is based on a combination of architectural and cytological changes [184], but evaluation of these changes is subjective and known to be inconsistent due to considerable inter- and intra-observer variations in the grading of lesions [12]. Digital pathology, which leverages the power of whole slide imaging and computer-aided diagnosis, holds great promise to providing rapid, consistent, and quantitative cancer diagnosis from histopathology images. Furthermore, noninvasive alternatives, such as various kinds of optical imaging techniques, have been sought to avoid the pain and discomfort of the biopsy procedures. Recent advancements in hyperspectral cameras, image analysis methods, and computational power have led to the development of

hyperspectral imaging (HSI) system as a promising diagnostic tool for early cancer detection [189].

HSI is an emerging optical modality that combines spectroscopy and wide-field imaging, which can rapidly interrogate large tissue surfaces and do not require tissue removal. Although spectroscopy has been explored extensively for probing molecular, cellular, and tissue properties and characterizing correlation of tissue parameters with disease state, such fundamental research has not been investigated vigorously in HSI. Many existing studies have been focused on developing new hardware systems. Very few efforts have been dedicated to investigate the underlying biological rationale of HSI for cancer detection, and the biological origins of differences in the measured reflectance signals of normal and neoplastic tissue are not well understood.

HSI provides an indirect measurement of the underlying tissue biochemical and morphological properties. The spectral characteristics of diffuse reflectance from heterogeneous biological tissue are the result of a complex interplay of the intrinsic absorption and scattering properties of the tissue, the distribution of chromophores and scatterers, together with the source-tissue-detector geometry [190]. Thus, the biochemical and/or structural characteristics of the biological tissue determine the intrinsic absorption and scattering properties of the tissues, which in turn generate the measured reflectance signature. In the field of diffuse reflectance spectroscopy, researchers have reported a variety of modeling methods to inversely estimate the absorption and scattering coefficients μ_a and μ_s, respectively, from the diffuse reflectance to characterize tissue properties. These modeling methods provide a way to connect the spectral features with the underlying biochemistry and morphology. However, they generally rely on the assumption of simplified tissue composition and structure, and specific source-detector settings.

To meaningfully interpret the hyperspectral dataset, there is a requirement to relate tissue architecture and morphology that occur with neoplasia to the bulk optical signal measured. On a microscopic scale, the architectural and cytological changes in neoplastic tissue can be quantified by a variety of histological features extracted from digitized pathological images for computer-aided diagnosis. Meanwhile, the spectral signature at each image point reflects the macroscopic features of corresponding tissue. Thus, we hypothesize that the spectral signature measured by HSI has significant association with histological features which quantify the tissue architectural and morphological alterations during neoplastic transformation. To validate this hypothesis, a predictive model is developed to combine multiple pathological features, including color, texture, morphometry, and topology features from epithelium tissue and its constituent nuclei and cytoplasm, for computer-aided diagnosis of tongue neoplasia. Next, an optimal feature subset is selected from the list of original histological features to best distinguish normal tissue from neoplastic tissue in histological images. Finally, the correlation coefficients between the spectral signature of both *in vivo* and fresh *ex vivo* tongues and the optimal histological features of the corresponding histological images are calculated.

4.2.2 Dataset

The H&E slides of tongue tissues from the 4NQO-induced carcinogenesis model are digitized and reviewed by an experienced pathologist specialized in head and neck cancer. The tongue dorsal surface on each whole-slide pathological image is segmented into pathological regions of normal, dysplasia, carcinoma in situ (CIS), and squamous cell carcinoma. We crop 1472 × 922 images from the whole-slide pathological image at 20× magnification using the Aperio ImageScope software (Leica Biosystems). A total of 1157 images are generated for quantitative analysis as listed in Table 7.

Table 7: Summary of Pathological Images from a Mouse Tongue Carcinogenesis Model

Mouse ID	Normal	Dysplasia	CIS	Carcinoma	Total Number of Images
M1	45	68	25	20	158
M2	79	56	51	4	190
M3	49	51	21	0	121
M4	18	40	9	0	67
M5	80	92	10	0	182
M6	26	38	8	0	72
M7	29	56	0	0	85
M8	70	53	0	0	123
M9	13	52	0	0	65
M10	13	81	0	0	94
Total Number of Images	422	587	124	24	1157

4.2.3 Method Overview

Figure 27 shows the method flowchart for correlation analysis between spectral signatures and histological features. The first step is to map the pathology gold standard onto the dorsal surface of the tongue hypercube. Each line on the color-coded pathology map may consist of multiple pathology types. The reflectance spectra of individual pixels within each pathological category along each line are averaged to generate the spectral signature of healthy, dysplasia, carcinoma in situ, and carcinoma tissue. Each vertical line on the pathology map corresponds to a whole-slide histological image of the tongue. The next step is pathological feature mining. Multiple sub-images within each pathology category are cropped and preprocessed for feature extraction and selection. An optimal feature subset that best distinguishes between normal and neoplastic tissue is identified. Finally, the correlation coefficients between the spectral signature and the most distinctive pathological features from the same tongue regions are calculated and displayed in a heatmap.

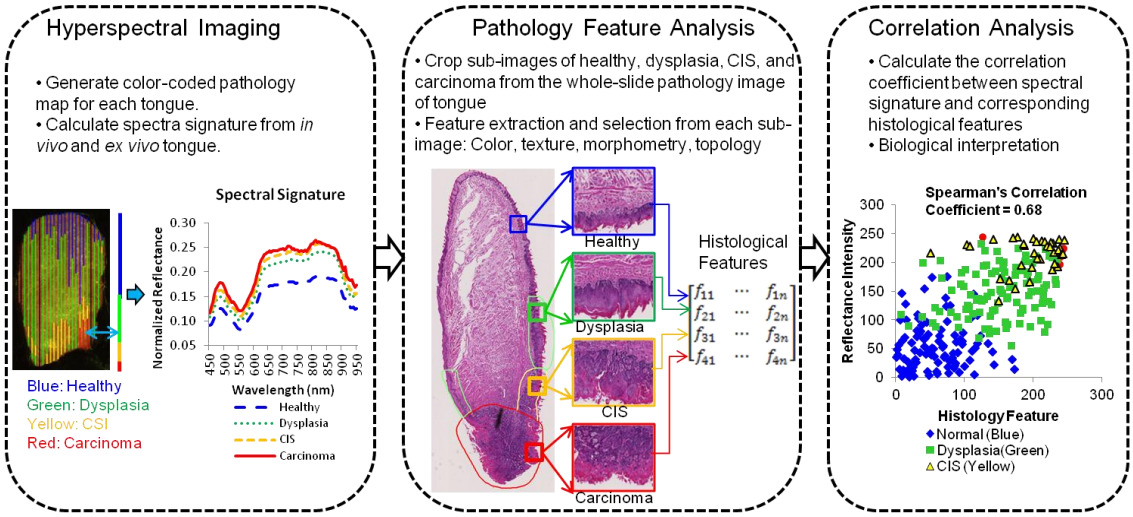


Figure 27: Flowchart for correlation analysis between spectral signature and histological features

4.2.4 Pathology Feature Mining

Figure 28 illustrates the steps for pathology feature mining. Since the focus of this thesis is about hyperspectral imaging, the method for histological image processing will only be briefly summarized in this chapter.

First, epithelium tissue was segmented from the connective tissue, muscle, and background. In clinical practices, the diagnosis of precursor lesions is based on the altered epithelium with an increased likelihood for progression to squamous cell carcinoma [191]. This step preprocesses the histological images for quantitative diagnosis.

Second, the epithelium image generated from the first step was further segmented into nuclei, cytoplasm, and background. The colors consisting of blue-purple, pink, and white in the histological image allowed a clear distinguishing between different cellular components within the epithelium. Therefore, k-means clustering using the Euclidean distance was applied on RGB color image of the epithelial layer to segment the epithelium into its three constituent components: nuclei, cytoplasm, and background. The

initial nuclear masks generated by k-means clustering suffered from the problems of small spurious background noise, small holes within the nuclei region due to nuclear inhomogeneity, and nuclear clusters with overlapping nuclei. To address the first two problems, a series of morphological operations were conducted. To separate touching nuclei, repeated K-means clustering was used to break down large nuclear clusters into small nuclear clusters and then marker-controlled watershed segmentation was employed to segment small nuclear clusters into individual nuclei. Ellipse fitting was finally conducted to get a smooth contour of individual nuclei.

Next, multiple features were extracted from both the entire epithelial image and its constituent components including nuclei and cytoplasm. To quantify the abnormal tissue changes during carcinogenesis, 71 color and 149 texture features were extracted from each epithelium image and each cytoplasm image, respectively. To quantify these changes, 71 color, 149 texture, 44 morphometric, and 8 topological features were extracted from each nuclei image. The color and texture features were the same as extracted from the epithelium and cytoplasm image. In total, we extracted 712 features from each epithelium image and its constituent nuclei and cytoplasm images. Table 8 is a summary of all the features.

Finally, feature selection and supervised classification with nested cross validation was conducted to build predictive models for cancer diagnosis. The image set used for normal/neoplastic classification consisted of a total of 1157 images (735 neoplastic and 422 normal images) from 10 mouse tongues. To identify a compact and distinctive feature subset from the 721 features of each histology image, we performed feature selection with the mRMR framework described in Chapter 5, and built predictive models through nested cross validation (CV) consisting of leave-one-out outer CV and leave-one-out inner CV. The outer CV loop was used to estimate the classification performance and the inner CV loop was used to tune the optimal parameters for the

model development. Each run of the ten-fold outer CV algorithm consisted of training models on image sets from nine tongues and testing on image sets from the remaining tongue. A nine-fold inner CV was conducted to select the optimal feature numbers from the subsets of the training data from eight tongues and to validate the model using the remaining subset. Support vector machine with a Gaussain Radial Basis function (RBF) was used as the classifier. Parameters were optimized via grid search over a pre-defined range. We considered the feature dimension m over the range of [1,5,10,30,50,70,90,100,200,300,342], and the cost values $c \in 2^{[-5,-3,-1,1,3,5]}$ and kernel parameters $\gamma \epsilon 2^{[-5,-3,-1,1,3,5]}$. We assessed the performance of the classifiers with accuracy, sensitivity, and specificity.

Table 8: Summary of All the Histological Features

Tissue Components	Feature Name	Feature Type	Feature Number
Epithelium	***Color***: Transformed RGB histogram; Red-blue difference histogram; Statistical measures (mean, median, standard deviation, minimum, maximum, skewness, and kurtosis) of the red-blue intensity difference.	color, texture	220
	Texture: Haralick feature, Local binary pattern, Fractal textures, Gabor filter (energy, entropy)		
Cytoplasm	***Color*** feature same as those extracted from Epithelium.	color, texture	220
	Texture feature same as those extracted from Epithelium.		
Nuclei	***Color*** feature same as those extracted from Epithelium.	color, texture, morphometry, topology	272
	Texture feature same as those extracted from Epithelium..		
	Morphometry: Statistical measures (mean, median, standard deviation, minimum, maximum, skewness, and kurtosis) of the nuclear size, solidity, eccentricity, major axis length, minor axis length, compactness, neighborhood radius; Nuclei to cytoplasm ratio;		
	Topology: Statistical measure (mean, maximum, minimum, disorder) of the area and perimeter of the Delaunay triangulation		

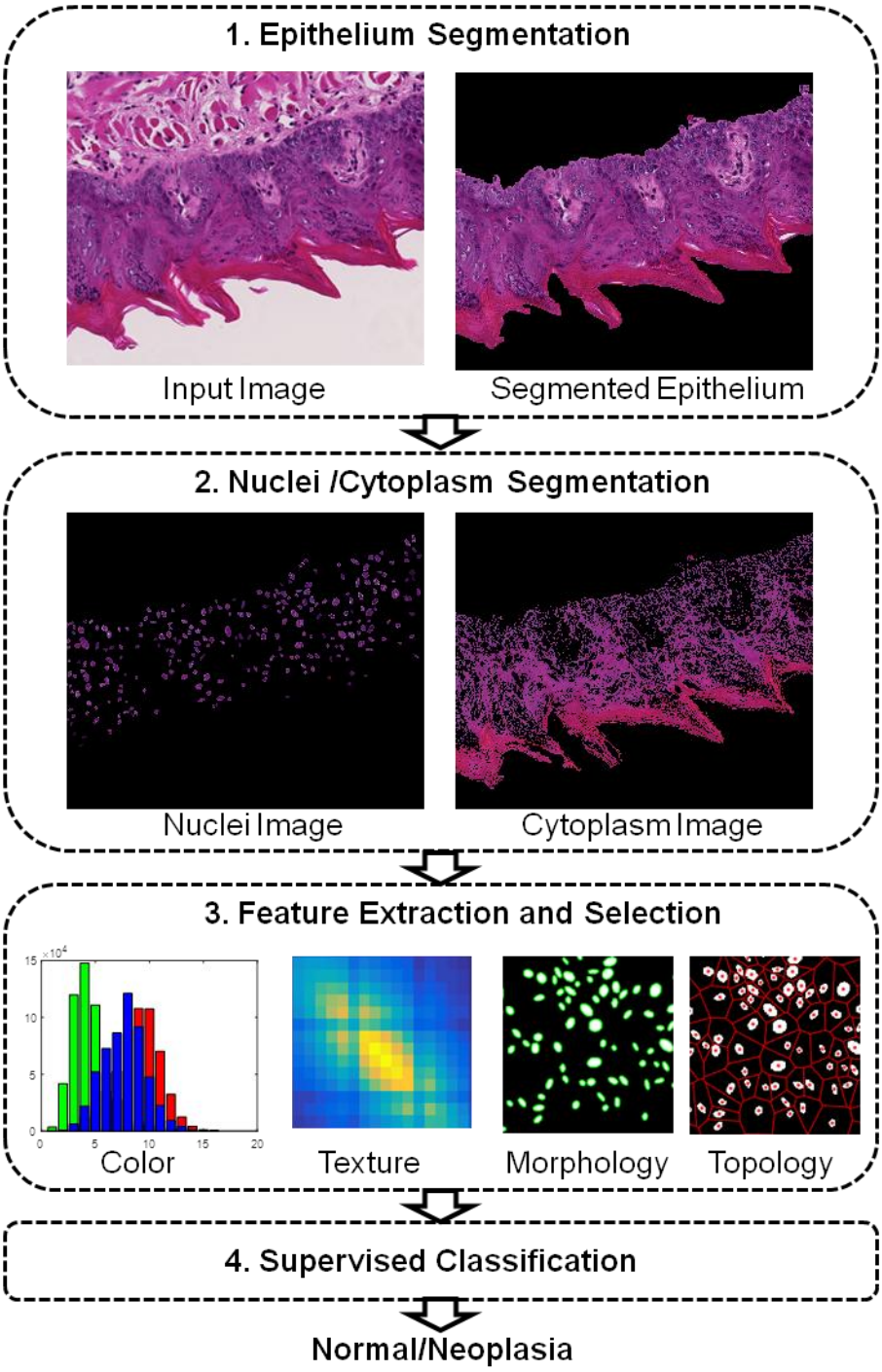


Figure 28: Flowchart for pathological feature mining.

4.2.5 Correlation between Spectral Signature and Histology Features

Spearman's rank correlation coefficient is a nonparametric approach for evaluating the degree of monotonic association or correlation between two independent variables [192]. We use Spearman's rank correlation coefficient to assess the association between spectral signatures and the selected optimal histology feature subset from all pathological tissue types. To test the significance of each pairwise correlation, we assume that there is no correlation between the two variables and reject the null hypothesis when the p-value is less than or equal to the significance level α=0.05. The resulting matrix of pairwise correlation coefficients is visualized as a heatmap where the positive and negative correlation coefficients are displayed in green and red, respectively. Correlation coefficients that are not statistically significant are displayed in white. All statistical analysis was performed in Matlab R2015b.

4.2.6 Results

Spectral Signature

Figure 29 showed representative spectral signatures of healthy and neoplastic tissue (dysplasia/CIS/carcinoma) measured by HSI from *in vivo* and *ex vivo* mouse tongues. The overall shapes of these spectra appeared similar for all the pathology types, which were generally smooth and broad in the visible and near-infrared region. The characteristic dips around 540 nm coincided with hemoglobin's absorption peaks. The reflectance intensity was generally weaker for healthy tissues than for neoplastic tissue including dysplasia, CIS, and carcinoma in both *in vivo* and *ex vivo* mouse tongues.

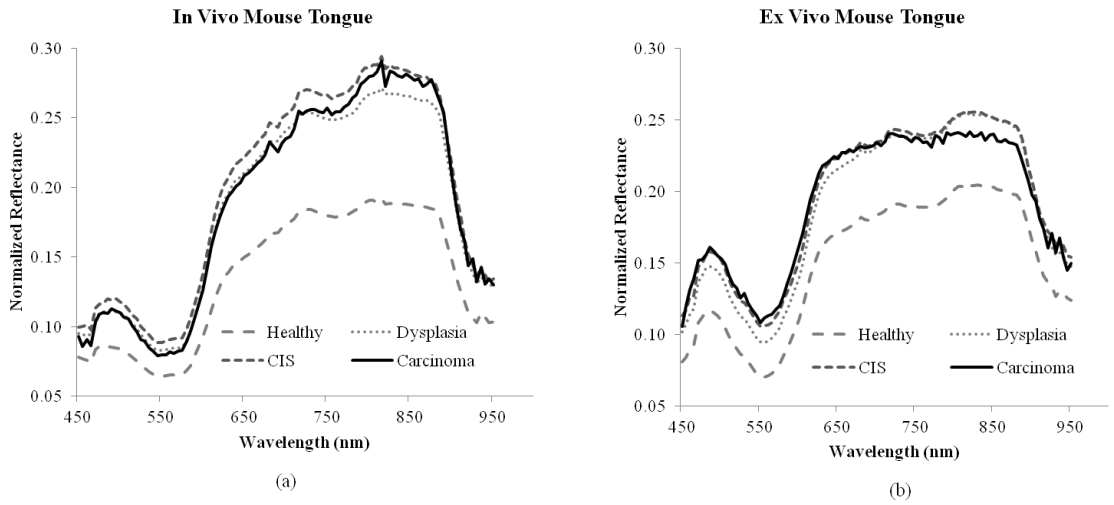


Figure 29: Reflectance spectral signature from hyperspectral images of (a) *in vivo* mouse tongue and (b) *ex vivo* mouse tongue.

Feature Selection

Figure 30 (a) plotted the CV accuracy as a function of feature dimension. We can see that when the number of features reached 30, classification accuracy reached the maximum. Since mRMR method worked by incrementally adding features according to maximal relevance and minimal redundancy criterion, we can further look at which feature was the most frequently selected to be the first, second, up until the 30th feature during the cross validation. In Figure 30 (b), the horizontal axis was the feature rank, and the vertical axis was feature categories. Bright color means the feature is highly selected. Figure 30 (b) illustrated that among the best 30 features, texture features extracted from epithelium and its constituent components: cytoplasm and nuclei were the most frequently selected feature type across all the 30 ranks for the distinction of normal tissue from neoplasia. Mathematically, texture features characterize differences in the spatial arrangement of gray values of neighboring pixels, which have been shown to be effective at quantifying the tissue structural changes for oral cancer grading in [193].

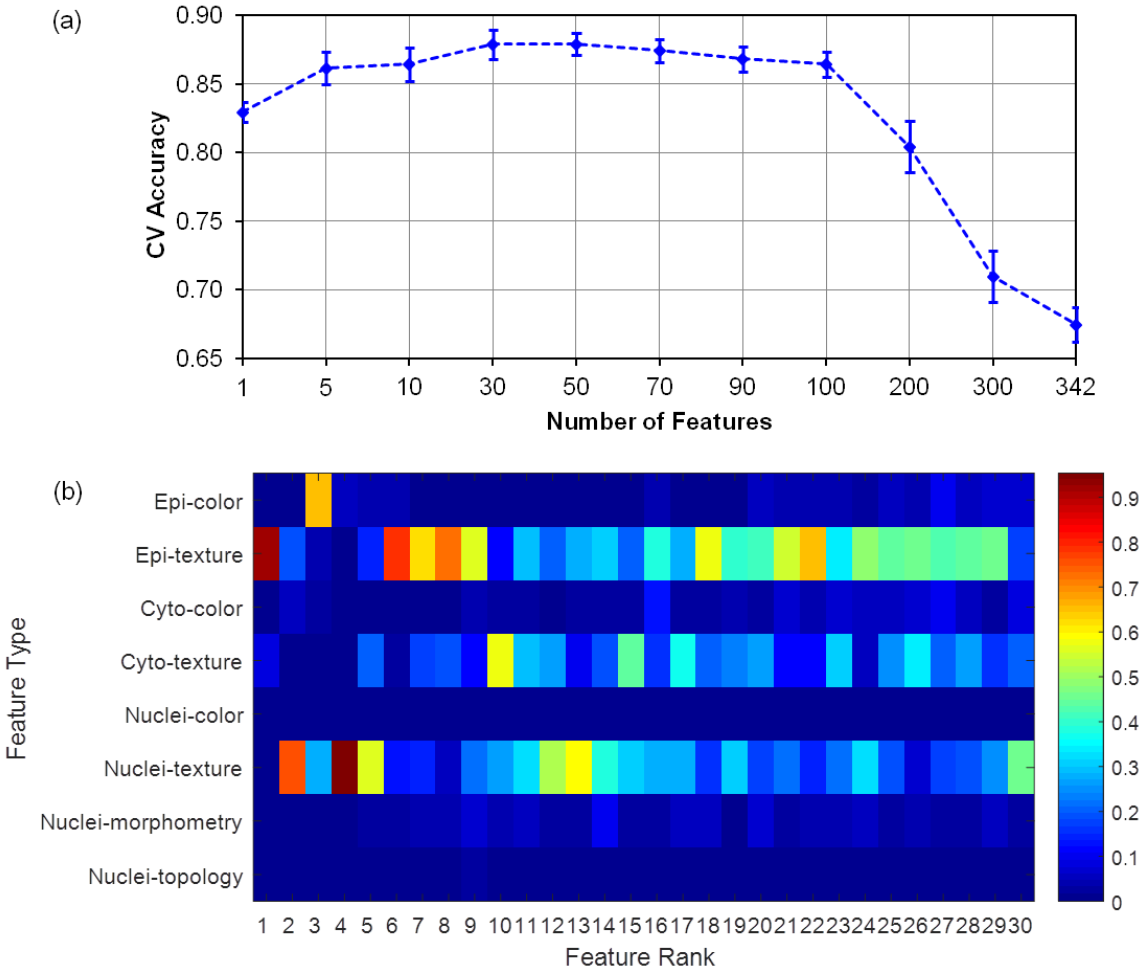


Figure 30: Feature selection and ranking frequency

Predictive Analysis

To evaluate the robustness of our model, Figure 31 (a) showed an analysis of the predictivity of selected model parameters, i.e., the correlation of model performance in internal and external validation. Figure 31 (b) showed the confusion matrix for neoplasia detection, which had a sensitivity of 92.7% and a specificity of 82.7%.

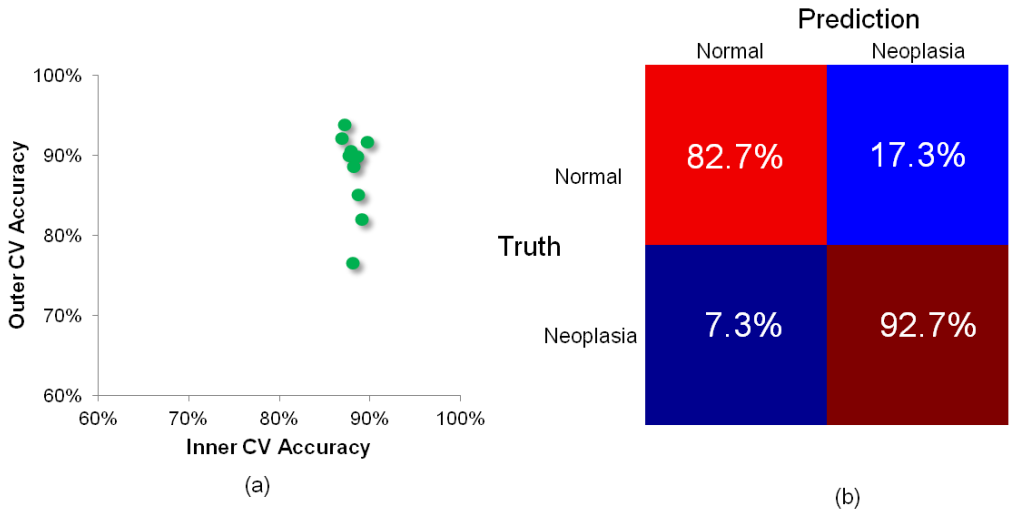


Figure 31: Predictive modeling for the distinction of neoplasia from non-neoplastic tissue. (a) Predictivity analysis of the selected models with the optimal parameters via grid search. (b) Confusion matrix of labeling on outer cross validation.

Correlation Coefficients

Figure 32 shows the correlation heatmap of pairwise association between the optimal histological feature subsets which are found to have the best distinguishing power for neoplasia detection and spectral signature from hyperspectral imagers of *in vivo* and *ex vivo* tongues. The average and standard deviation of the correlation coefficients for each row are listed next to the corresponding heatmap. We have three major observations from the heatmaps. Firstly, spectral signature from HSI is significantly associated with the optimal histological feature set, which suggests that HSI captures the key diagnostic information reflecting the tissue architectural and morphological changes during neoplastic transformation. Secondly, the strengths of the correlation between each histology feature and reflectance intensities over all the wavelengths are very similar, which suggests that the diagnostic importance of the spectral information for cancer detection. Lastly, comparing the average Spearman's

correlation coefficients of each histology feature with wavelengths, the spectral signature of *in vivo* tongues exhibits stronger association with histological features than that of *ex vivo* tongues, which suggests that hyperspectral images of *in vivo* tongues capture more diagnostic information than those of *ex vivo* tongues.

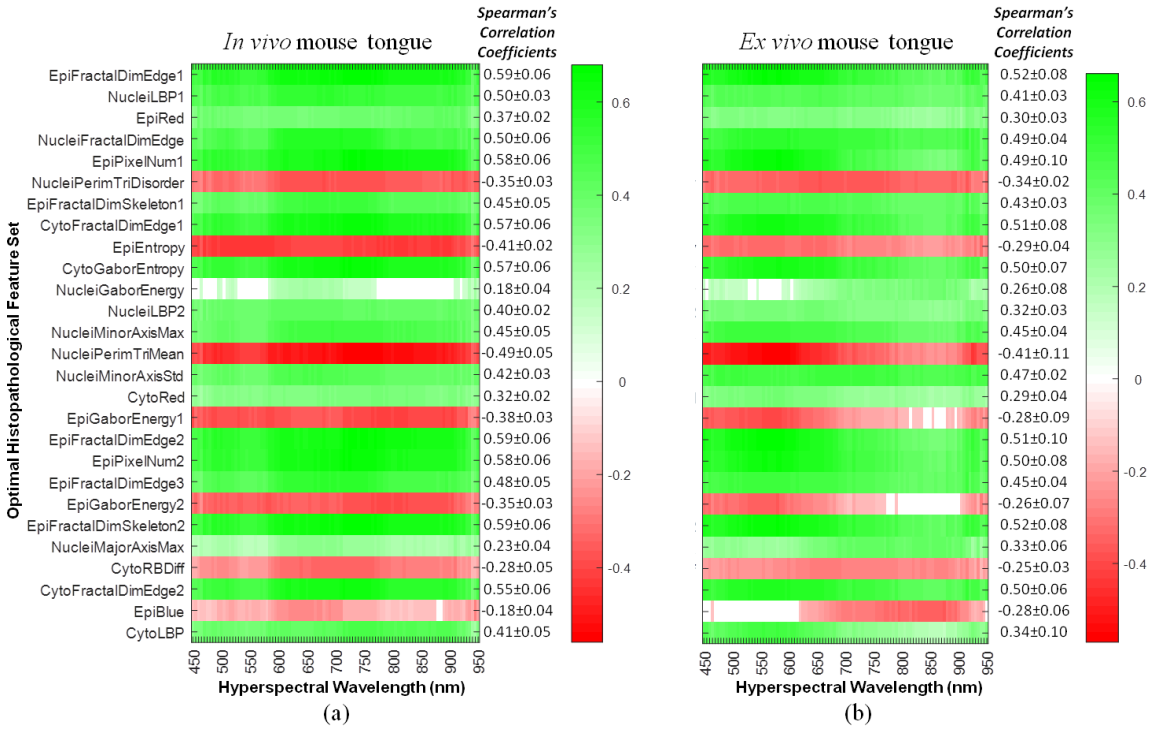


Figure 32: Correlation heatmap showing Spearman's correlation coefficient between spectral signature (horizontal axis) and the selected optimal histology feature subset (vertical). Green = positive correlation, red = negative correlation, white = no correlation or correlations that are not statistically significant.

Table 9 summarizes all the nine histology features which are highly and significantly associated with spectral signatures (Average correlation coefficients $r_s \geq 0.5$ or $r_s \leq -0.5$). Seven out of the nine features are quantifying the textural changes in tissue. Fractal dimensions extracted from the epithelium, nuclei, and cytoplasm are the most frequently selected features with strong and significant

correlation with spectral signature. Only one histology feature has strong enough negative correlation: the mean perimeter of the Delaunay Triangulation constructed from the segmented nuclei image.

Table 9: Summary of Representative Histological Features ($\|r_s\| \geq 0.5$)

Location	Feature Name	r_s	Feature Explanation
Epithelium	Fractal dimension (edge)	0.59	Quantitative description of complex, irregularly shaped objects in epithelium
	Fractal dimension (skeleton)	0.59	
	Pixel number (edge)	0.58	
Nuclei	Fractal dimension (edge)	0.59	Quantitative description of complex, irregularly shaped nuclear objects
	Local binary pattern	0.5	Rotation-invariant texture feature characterizing spatial structure and contrast of nuclei
	Minor axis length (max)	0.45	Quantify the variations in nuclear size and shape
Cytoplasm	Fractal dimension (edge)	0.57	Quantitative description of complex, irregularly shapes in cytoplasm
	Gabor texture (entropy)	0.57	Characterize the randomness in texture of Gabor magnitude of cytoplasm image
Nuclei	Perimeter of Delaunay triangulation (mean)	-0.49	Describe the distances between individual nuclei

Furthermore, we plot the distribution of two representative histology features with significant and strong positive and negative correlation with spectral signatures as shown in Figure 33 (a) and (c). As the degree of tissue malignancy increases, the fractal dimension of the epithelium tends to increase (Figure 33 (a)), which reflects the abnormal structural and morphological changes of tissue during neoplastic transformation, such as the loss of cellular organization and the increase in the number of atypical nuclei [184]. Meanwhile, as the fractal dimension of the epithelium increases, the reflectance intensities also tends to increase as shown in Figure 33 (b), which shows positive correlation between histology feature and spectral signature. On the other hand, the mean

perimeter of the Delaunay triangles tends to decrease as tissues transform from benign to malignant (Figure 33 (c)), which is consistent with the fact that nuclei have proliferated and become more crowded in malignant tissue. As shown in Figure 33 (d), the reflectance intensity tends to increase as the mean perimeter of Delaunay triangles decreases, exhibiting negative association between the two.

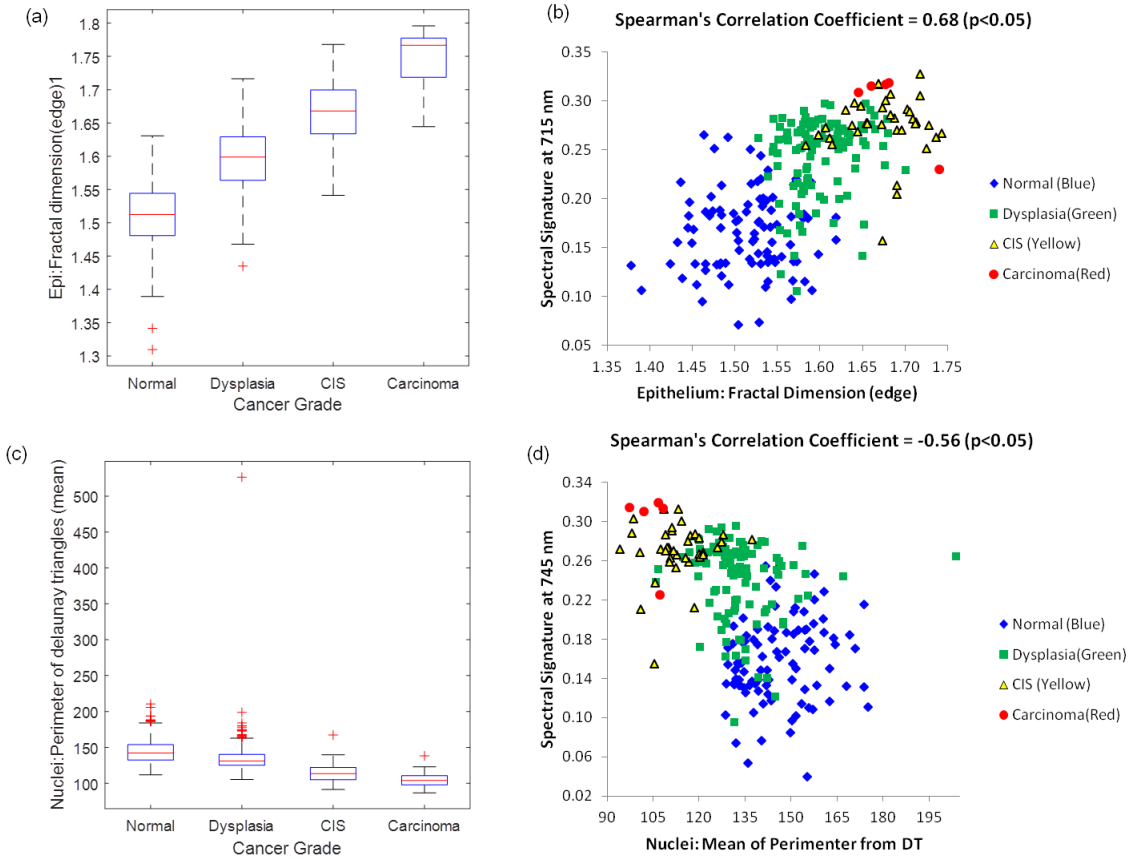


Figure 33: Distribution of histological features and corresponding scatter plots with spectral signature at selected wavelengths. In (a) and (b), the histology feature is fractal dimension extracted from epithelium, which has strong and significant correlation coefficients with spectral signature at 715 nm. In (c) and (d), the histology feature is the mean perimeter of Delaunay triangles extracted from nuclei image, which exhibits significant negative correlation with spectral signature at 745 nm.

4.2.7 Discussion

The major contribution of this study is that the successful validation of the hypothesis that the spectral signature has significant association with histology features that reflects the tissue architectural changes during malignant transformation. Multiple quantitative histological features were extracted and selected from the epithelium and its constituent components-nuclei and cytoplasm, which best distinguish normal from neoplastic tissue. The selected optimal feature subset is a combination of color, texture, morphometry, and topology features, which quantify the above architectural and morphological changes during tissue malignant transformation. We note that the average reflectance intensities of neoplasia are stronger than that of normal tissue in both *in vivo* and *ex vivo* mouse tongues. This observation is consistent with the representative *in vivo* reflectance spectra of nonmalignant and malignant tissue from patients with head and neck cancers [194], and is also consistent with the *ex vivo* reflectance spectra of malignant and adjoining normal tissue from the tongue cancer patients [188]. Furthermore, we also observe for the first time that reflectance intensity is increasing with increased fractal dimension and other texture features throughout neoplastic progression. The increase of fractal dimension is associated with the structural changes in epithelium, such as increased nuclear size, atypical nuclear shape, increased DNA content, and hyperchromasia with coarse and irregular chromatin clumping. One possible explanation for the spectral difference between normal and neoplastic tongue tissue can be that light scattering events inside the epithelium tissues change significantly with the progressive development of squamous lesions, thus leading to the alterations in the diffuse reflectance spectrum and forming a potential physiologic basis for cancer detection with hyperspectral imaging. More specifically, the scattering coefficient μ_s (in units of mm^{-1}) is defined as the rates of radiant energy loss due to scatter per incremental unit photon path length in a tissue. Qualitatively, for a given tissue volume and fixed absorption conditions, the diffuse reflectance increases as tissue scattering μ_s increases

[190]. Therefore, the alteration of the scattering density and distribution in the epithelium is likely to be contributing to the increased reflectance in neoplastic tissue. The scattering is relatively homogeneous across all the wavelengths, which may explain the homogeneous association between spectral signature and individual histological features.

To better understand the scattering origin in tongue, we further look into the histological structures of the tongue. As shown in Figure 34, mouse tongues have layered structures similar to human tongues, which include an outer stratified squamous epithelium, overlying and attached to a dense connective tissue called the lamina propria, and deep skeletal muscle fibers. According to [195], the epithelium consists of four cellular layers with tightly packed epithelial cells with varying degree of differentiation, beginning with the basal layer of undifferentiated cells that divide continuously, through layers of suprabasal cells into mature cells in the outermost keratinized layer. The tissue components in lamina propria mainly include fibroblasts, collagen fibers, and blood vessels in the form of capillaries. The skeletal muscle tissue mainly includes bundles of striated muscles with many blood vessels and nerves between them.

In the therapeutic window from about 600 to 1300 nm, biological tissue have relatively low absorption and high, forward-directed scatter, which allows both substantial penetration of light into tissue and high remittance of light scattered out of the tissue after deep penetration [190]. Light is scattered whenever it encounters refractive index variations. Tissue structures such as membranes, nuclei, mitochondria, and other organelles all have different refractive indices from the surrounding cytoplasm, making tissue a highly-scattering medium. In addition, the extracellular matrix containing keratin and collagen fibers are highly scattering. The total signal reflected from tissue can be divided into two parts: single backscattering from the uppermost tissue structures such as cell nuclei and multiple scattering from deeper tissue layers. The epithelium mainly contributes to the spectral response of photons experiencing few scattering events. The deeper tissue structures contribute to multiple scattering events. As tissue transform from

benign to malignant tumors, a series of cytological and architectural changes occur progressively, which in turn alters the distribution, density, size, and shape of major light scatters such as collagen, keratin, cell nuclei, mitochondria, and other cytoplasmic organelles during cancer progression. All these architectural and cytological changes could affect the scattering properties of tissue, which in turn change the reflectance signal measured by HSI. A key morphological change in dysplasia is the loss of stratification due to a lack of normal maturation from the cells in basement membrane to the surface keratin. Cells located in the surface layer of the epithelium have the same immature appearance as those in the deep basal layers [196]. This study suggests that the structural changes of the neoplastic tongue such as increased epithelium thickness and increased nuclear density may increase the backscattering of light, and the increased scattering may contribute to the spectral difference between neoplastic and normal tissue. These hypotheses remain to be investigated by future studies.

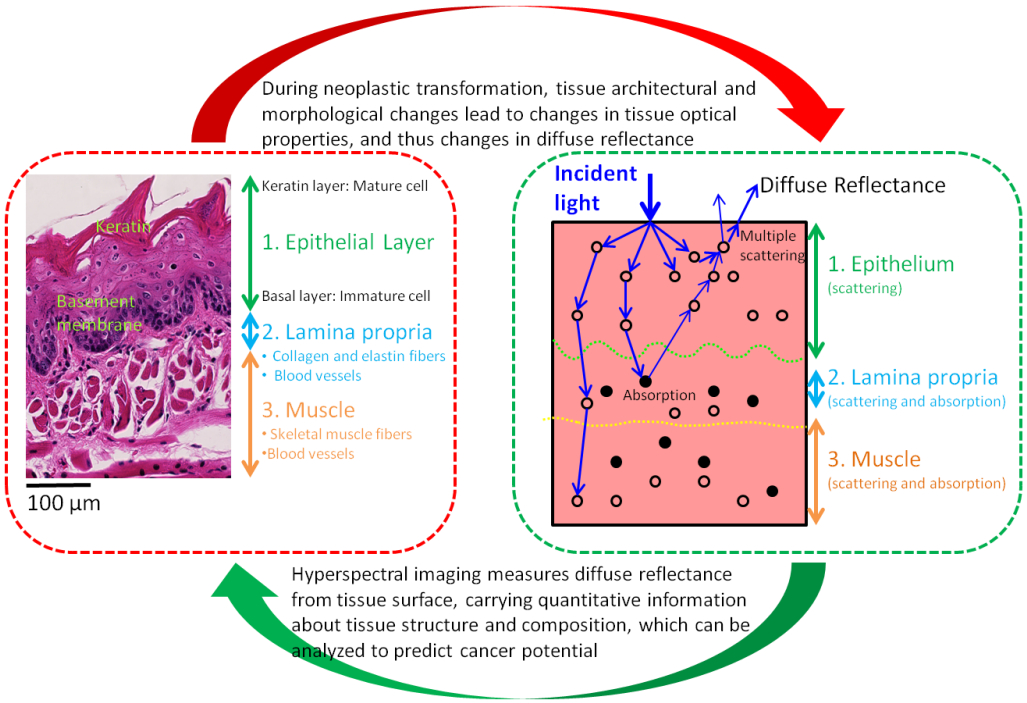


Figure 34: Interpretation of the association between HSI and histological features.

4.2.8 Conclusion

In summary, we have demonstrated that the spectral difference between normal and neoplastic tissue of mouse tongues, which are measured by HSI in a macroscopic scale, is associated with the histological features that can quantify the tissue architectural alterations during neoplastic transformation. This study provided some insights into the interpretation of cancer data acquired with hyperspectral imaging.

CHAPTER 5

INTRAOPERATIVE CANCER DETECTION WITH HYPERSPECTRAL IMAGING AND IMAGE QUANTIFICATION IN AN ANIMAL TUMOR SURGERY

Chapter 3 and chapter 4 have reported the use of HSI as a noninvasive diagnostic tool for the detection and delineation of head and neck cancer in a subcutaneous xenograft tumor model and a chemically-induced tongue carcinogenesis model. This chapter aims to investigate the potential of HSI as a diagnostic tool for rapid cancer detection during image-guided surgery. The data from hyperspectral imaging often needs to be processed appropriately in order to extract the maximum useful information that differentiates cancer from normal tissue. In this chapter, we propose a framework for hyperspectral image processing and quantification, which includes a set of steps including image preprocessing, glare removal, feature extraction, and ultimately image classification. The framework has been tested on images from mice with head and neck cancer using spectra with wavelengths from 450 to 900 nm. The image analysis computed Fourier coefficients, normalized reflectance, mean, and spectral derivatives for improved accuracy. The experimental results demonstrated the feasibility of the hyperspectral image processing and quantification framework for cancer detection during the animal tumor surgery.

5.1 Introduction

Surgery remains the most definitive treatment for most solid tumors [15]. During surgery, complete tumor removal is essential for the postoperative prognosis of patients [197]. Positive resection margins could lead to local recurrence of cancer [198], additional surgeries [199], and increased mortality for cancer patients [200] [201]. Thus,

improvement in the completeness of tumor removal would benefit patients and might produce significant cost savings [15].

During tumor surgery, it can be challenging to differentiate cancer from normal tissue during the process of resection. Visual inspection and palpation are routinely used to determine tumor margin during surgery [16]. This process can be subjective and inconclusive. Intraoperative frozen tissue evaluation is also use to estimate surgical margin at initial surgery [15], which may suffer from errors that occur during sampling and histological interpretation, and pressure is always mounting to reduce this time and cost factor. In addition, the histological processing can take time [17] and extends anesthesia-related risks. Therefore, there is a clinical need to develop real-time imaging methods to help the surgeon for localization and assessment of tumor margins during surgery.

Hyperspectral image-guided surgery has been reported in the literature. Panasyuk *et al.* [60] utilized HSI to detected small residual tumors of 0.5 to 1 mm in a breast cancer surgery of a rodent tumor model. Kiyotoki *et al.* [202] reported the use of HSI technology for the distinction of gastric cancer from normal mucosa in endoscopically resected lesions. Gebhart *et al*. [203] used an imaging system that combined fluorescence and diffuse reflectance imaging for human brain tumor resection guidance. In addition, HSI has also been utilized for differentiating key anatomic structures during surgery such as differentiation of the bile duct from surrounding arteries [204]. A more comprehensive summary about surgical guidance with HSI can be found in our recent review paper [189]. Although preliminary studies have demonstrated the feasibility of HSI towards an intraoperative visual aid, the diagnostic capability of HSI is largely dependent on the interpretation and development of hyperspectral image analysis methods. Previous efforts have been mainly focused on the design and test of various HSI systems while the pre-processing and analysis of surgical hyperspectral images are not well investigated. The challenges in developing a diagnostic support system for rapid cancer detection mainly

lie in two aspects. Firstly, the intraoperative hyperspectral images are usually distorted due to motion artifacts, glare (also known as specular reflection), and curvature variations, which require rigorous pre-processing procedures before further analysis. Secondly, the analysis of hyperspectral data is time-consuming due to the large volume and high dimensionality. On the other hand, real-time surgical guidance requires the analysis to be complete within a certain time frame. So it is essential to extract and select the most relevant features to reduce dimensionality without sacrificing diagnostic accuracy.

The objective of this study was to (1) develop a set of pre-processing techniques to reduce intraoperative image distortions; (2) conduct feature extraction and selection to identify the most relevant spectral features and reduce data dimensionality; and (3) evaluate the diagnostic potential of the selected feature set for cancer detection. The major contribution is the development of a comprehensive framework for surgical hyperspectral image processing and analysis which could facilitate the advancement of medical hyperspectral imaging towards clinical translation. The rest of the chapter is organized as follows. Section 5.2 describes the details about hyperspectral imaging instrument, animal experiment, and data acquisition. Section 5.3 explains the proposed image analysis framework, including pre-processing, feature extraction, feature selection, and classification techniques. Section 5.4 presents the analysis and validation results. Section 5 discusses the contribution, limitation, and future development of the proposed methods. Finally, conclusions are drawn in Section 6.

5.2 Materials and Experimental Design

5.2.1 Hyperspectral Imaging System

The HSI system used in this study is the same as that used in previous chapters.

5.2.2 Animal Imaging and Tumor Surgical Experiments

A head and neck tumor xenograft model using head and neck squamous cell carcinoma (HNSCC) cell line M4E with green fluorescence protein (GFP) [205] was adopted in the experiment. The HNSCC M4E cells were maintained as a monolayer culture in Dulbecco's modified Eagle's medium (DMEM)/F12 medium (1:1) supplemented with 10% fetal bovine serum (FBS) [14]. M4E-GFP cells which are generated by transfection of a pLVTHM vector into M4E cells were maintained in the same condition as M4E cells. The animal experiment was approved by the Animal Care and Use Committee (IACUC) of Emory University.

In this experiment, 8 female nude mice aged 4-6 weeks were injected with 2×10^6 M4E cells with GFP on the lower back. Surgery was performed approximately one month after tumor cell injection. Before surgery, mice were anesthetized with a continuous supply of 2% isoflurane in oxygen. After the anesthesia administration, the skin covering the tumor was removed to expose the tumor to simulate a surgical situation.

Reflectance Image Acquisition: Hyperspectral images with the interior infrared, the white excitation, and an auto setting for exposure time, were captured over the exposed tumor. Each hypercube contain 226 spectral bands from 450 nm to 900 nm with 2 nm increments.

Fluorescence Image Acquisition: Subsequently, a 450 nm excitation filter and auto-exposure time were selected for the blue fluorescence image acquisition. Tumors manifested themselves as green light in the fluorescence images due to GFP.

Surgical Removal of Tumors and Histology Processing: After imaging, the tumors were removed horizontally from the bottom using a blade and were inked with four different colors to represent the head, tail, left, and right orientations of the tumors in the mice. Histological samples were kept in formalin and sent for histological evaluation after 24 hours. Histological slides with H&E staining were digitized to provide histological assessment of tumor margins.

5.2.3 Evaluation Methods for Cancer Detection

In this study, GFP fluorescence images provide the *in vivo* gold standard for tumor margin delineation. Although the current gold standard for cancer diagnosis continues to be the histological assessment of H&E stained tissue, the *ex vivo* tissue specimen undergoes deformations including shrinkage, tearing, and distortion, which makes it difficult to align the *ex vivo* gold standard with *in vivo* tumor tissue. However, the *in vivo* GFP fluorescence images provided a much better alignment with hyperspectral reflectance images in the animal model, since they were acquired *in vivo* immediately after the acquisition of reflectance images for each mouse while the tumor position and shape in reflectance images are exactly the same as those in the fluorescence images. The tumor and surrounding normal tissue exhibited high contrast in the GFP images. Since histopathology is the gold standard for cancer diagnosis in clinics, we also acquired the H&E stained histological image as the *ex vivo* gold standard.

In this paper, tumor regions were identified manually on the GFP images and the classification results were then compared with the manual maps. Since human tissue does not contain GFP naturally, registration methods are desirable to align the *in vivo* hyperspectral images with *ex vivo* histological images as discussed in our previous publication [149] in order to move forward for future human studies.

5.3 Framework for Hyperspectral Image Processing and Quantification

Although hyperspectral image analysis methods have been developed for over three decades in the remote sensing area, it is still in the early stages in biomedical applications. With high dimensional datasets, it is not a trivial task to extract the most relevant information from the raw data and classify them into tumor and non-tumor tissue. Automated intraoperative cancer detection is especially challenging due to the motion artifacts and glares caused by specular reflection [206], and the high dimensionality of the dataset [207]. Therefore, we proposed a comprehensive workflow

as shown in Figure 35 that includes a set of pre-processing techniques as well as feature extraction and selection methods for intraoperative hyperspectral image analysis. The performance of the selected feature set was evaluated by supervised classification method.

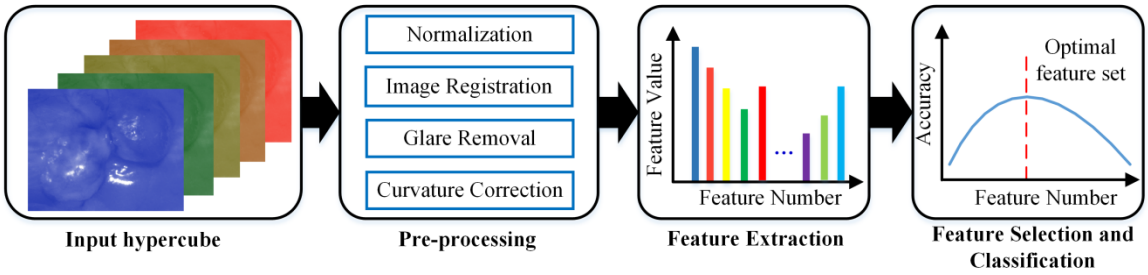


Figure 35: The flowchart of the proposed quantitative analysis framework.

5.3.1 Pre-processing methods for Hyperspectral Images

The pre-processing steps of the hypercube consist of data normalization, image registration, glare detection and removal, and curvature correction.

Reflectance Image Calibration

The purpose of data normalization was to remove the spectral non-uniformity of the illumination device and the influence of the dark current. The raw data was corrected by converting into relative reflectance data using Equation (1).

Registration of Hyperspectral Images for Motion Correction

The HSI instrument used in our study was a spectral-scanning system, which captured the whole scene with two-dimensional (2D) detector arrays in a single exposure and then stepped through wavelengths to complete the three-dimensional (3D) data cube.

The exposure time for each image band varied between 12~30 ms, so the acquisition of one hypercube took about 4 to 7 seconds. The breathing and heart motion of the mice during the image acquisition would lead to the misalignment of the images at different bands even though they are within the same hypercube, which would affect the shape and intensity of the spectral curve. Therefore, the distinction between cancer and normal tissue with pixel-wise classification approaches may be compromised, necessitating the need for image registration to align the image bands within each hypercube.

Intensity-based rigid registration was employed since the motion was mostly global movement such as translation. The sum of squared difference was chosen as the similarity metric since it reflected the movement of each pixel very well. Next, a reference image was selected and the geometric transformation would be applied to the other images so that they were aligned with the reference. Three methods were compared to choose the most suitable reference image: (1) register each image band to its neighboring image band; (2) register all image bands to the band with highest signal to noise ratio; (3) register all image bands to the average of all image bands. The third method was able to remove most of the movement within one hypercube, and outperformed the first two methods.

After the reflectance hypercube was registered, the corresponding fluorescence image at the GFP emission band also needed to be aligned to the reflectance hypercube. We registered the GFP image at the peak emission band to the mean of the reflectance hypercube and then manually outlined the tumor boarder in the registered GFP image. These tumor maps served as the gold standard for the supervised classification in this study.

Glare Removal Method for Hyperspectral Images

Glare, also called specular reflection, is the mirror-like reflection of incident light from a moist surface. Optical images such as those from an endoscope and colposcopy

images acquired during surgery are often strongly affected by glare spots in the images which present a major problem for surgical image analysis [206]. In hyperspectral imaging, glare alters the intensities of the pixels in each image band and consequently changes the spectral fingerprint which could introduce artifacts in feature extraction and hence deteriorate classification results. Therefore, glare pixels need to be detected and removed before feature extraction and classification.

Currently, no glare removal technique exists for intraoperative hyperspectral imaging. To better understand the difference of glare pixels, we compared the spectral fingerprint of glare pixels with non-glare pixels and identified the characteristics of glare pixels. It can be seen from Figure 36 that glare pixels not only had higher intensities but also showed larger variations along the spectral bands. Therefore, the first-order derivatives of the spectral curves of glare pixels were much larger than those of non-glare pixels as shown in Figure 36 (d).

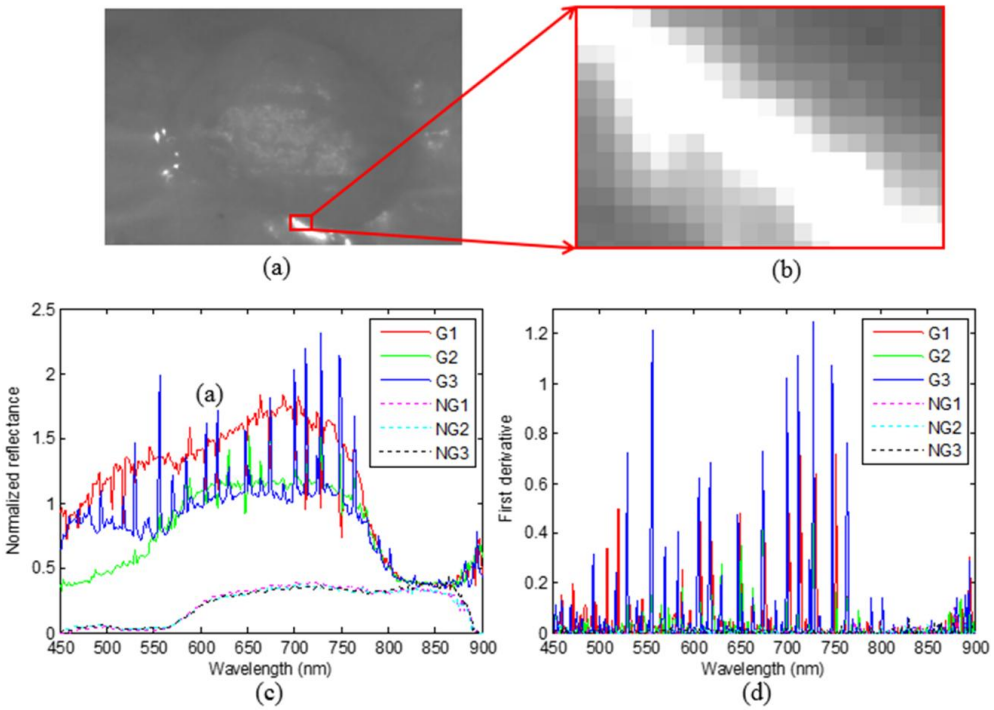


Figure 36: The rationale for the proposed glare detection method. (a) Image band at 758 nm. (b) Enlarged image of a selected glare region in the image band in (a). Glare pixels are much brighter than non-glare pixels. (c) The normalized reflectance curve of glare pixels (G1-G3) and non-glare pixels (NG1-NG3). The spectral curves of glare pixels vary significantly in many wavelengths. (d) The first-order derivative curves corresponding to the spectral curves shown in (c).

Based on the above observations, we proposed a three-step glare detection method: the first step was to estimate the first-order derivatives of spectral curves with a forward difference method. The second step was to calculate the standard deviation (SD) of each derivative curve and generate an SD image for each hypercube. Glare pixels show higher SD than normal pixels. Finally, compute the intensity histogram of each SD image, fit the histogram with a 'loglogistic' distribution (MATLAB function), and experimentally identify the threshold that separates glare and non-glare pixels.

Curvature Correction for Hyperspectral Images

Curvature correction method has been described in details in Chapter 2.

5.3.2 Feature Extraction from Hyperspectral Data

Feature extraction and representation are a crucial step for image classification tasks. Efficient feature extraction could lead to improved classification performance. However, hyperspectral imaging has only been recently applied to medical applications and it is not well understood which features are the most effective and efficient to differentiate cancer from normal tissue. Currently, one of the most frequently used features is the normalized reflectance curve of each pixel. With only the reflectance features, classification performance is still far from optimal. Therefore, there is still space for adding new features to improve the distinction of cancer from normal tissue.

In this study, we explored the utility of several spectral features which are derived from the reflectance curve of each pixel. Spatial features are not explored in this study but will be included in our future work. The extracted spectral features included: (1) first-order derivatives of each spectral curve which reflect the variations of spectral information across the wavelength range; (2) second-order derivatives of each spectral curve which reflect the concavity of the spectral curve; (3) mean, standard deviation, and total reflectance at each pixel which summarize the statistical characteristics of the spectral fingerprint; (4) Fourier coefficients (FC) which was initially found to be effective for target detection in the remote sensing field [208] and later was adopted for breast cancer margin classification from *ex-vivo* breast cancer hyperspectral images [209].

FC feature extraction involves transforming the original spectral $\mathrm{f(n)}, \mathrm{n} = 0,2, \ldots \mathrm{N}-1$ into the Fourier domain as $F(k) = \sum_{n=0}^{N-1} s(n) e^{-\frac{2\pi i}{N} kn}$ and then combining the selected real and imaginary components of Fourier coefficients $F(k)$ using the following formula:

$$\mathrm{FC(x)} = R(x) + I(x) \tag{7}$$

where

$$\mathrm{R(x)} = \begin{cases} Re\{F(x)\}, & if\ Re\{F(x)\} \geq Im\{F(x)\} \\ 0, & Otherwise \end{cases}$$

$$\mathrm{I(x)} = \begin{cases} Im\{F(x)\}, & if\ Im\{F(x)\} \geq Re\{F(x)\} \\ 0, & Otherwise \end{cases}$$

Different features may have very different numerical ranges, so each feature was standardized into its z-score (MATLAB function) by subtracting the mean from each feature and then dividing by its standard deviation.

5.3.3 Feature Selection Method

After feature extraction, the spectral feature dimension was increased from 226 to 904. Such a high spectral dimension poses significant challenges to the analysis of hypercubes. High dimensionality can significantly increase the computational burden and storage requirements, leading to increased data processing time, which is against the requirement of real-time tumor detection during surgery. Depending upon the wavelength range of imaging systems, different studies may have different reflectance features. It is not clear which wavelengths are most relevant in characterizing cancerous tissue and would provide a better contrast between cancer and normal tissue. In addition, with a narrow wavelength increment, there is likely spectral redundancy between adjacent bands. Last but not least, increasing the feature dimensionality without increasing the number of training samples may lead to a decrease in classification performance due to the curse of dimensionality i.e. the Hughes phenomenon [86]. Therefore, it is desirable to analyze the spectral redundancy in the high dimensional data and select the most characterizing compact feature set.

The goal of feature selection is to find a feature set *S* with n wavelengths $\{\lambda_i\}$, which "optimally" characterize the difference between cancerous and normal tissue. To achieve the "optimal" condition, we used the maximal relevance and minimal redundancy (mRMR) [210] framework to maximize the dependency of each spectral feature on the target class labels (tumor or normal), and minimize the redundancy among individual

features simultaneously. Relevance is characterized by mutual information I(x; y) which measures the level of similarity between two random variables x and y:

$$\mathrm{I}(\mathrm{x};\mathrm{y}) = \iint p(x,y)\log\frac{p(x,y)}{p(x)p(y)}\,dxdy \tag{8}$$

where $p(x,y)$ is the joint probability distribution function of x and y, and $p(x)$ and $p(y)$ are the marginal probability distribution functions of x and y, respectively.

We represent each pixel with M features $\Lambda = \{\lambda_i, i = 1, \ldots M\}$, $M = 904$, and the class label (tumor or normal) with c. Then the maximal relevance condition is to search features satisfying equation (3), which maximize the mean value of all mutual information values between individual features λ_i and class c:

$$\max D(s,c), \quad \mathrm{D} = \frac{1}{|S|}\sum_{\lambda_i \in S} I(\lambda_i, c) \tag{9}$$

The features selected by the maximal relevance condition are likely to have redundancy, which means that the dependency among these features could be large. When two features highly depend on each other, the respective class-discriminative power would not change much if one of them were removed. So the minimal redundancy condition can be added to select mutually exclusive features:

$$\min R(s), \quad R = \frac{1}{|S|^2}\sum_{\lambda_i,\lambda_j \in S} I(\lambda_i, \lambda_j) \tag{10}$$

The simple combination (equation (5) and (6)) of these two conditions forms the criterion "minimal-redundancy-maximal-relevance" (mRMR), which can optimize D and R simultaneously:

$$\max \Phi(D,R), \quad \Phi = D - R \tag{11}$$

In practice, incremental search methods can be used to find the near-optimal features defined by $\Phi(\cdot)$. Suppose we already identified a feature set S_{m-1} with $m-1$ features. The task is to select the m^{th} feature from the set $\{\Lambda - S_{m-1}\}$. This is done by

selecting the feature that maximizes $\Phi(\cdot)$. The respective incremental algorithm optimizes the following condition:

$$\max_{\lambda_j \in \Lambda - S_{m-1}} \left[I(\lambda_j; c) - \frac{1}{m-1} \sum_{\lambda_i \in S_{m-1}} I(\lambda_j; \lambda_i) \right] \quad (12)$$

5.3.4 Hyperspectral Image Classification

To determine the optimal feature set, the K-nearest neighbor (KNN) classifier was employed to evaluate the effectiveness of the selected features with leave-one-out cross validation. The workflow of the feature selection and classification was shown in Figure 37. First, eight hypercubes were separated into a training dataset with seven hypercubes and a testing dataset with the remaining hypercube. Next, 904 features were extracted from the training data and mRMR was used to select the optimal feature set on the training dataset as the feature number varied from 1 to 904. The selected feature set was applied to the testing dataset correspondingly. Next, KNN classifier was used to train the training data and to predict the labels of the testing data set with feature sets of different sizes. Finally, the feature set that gave the best classification performance was chosen as the optimal feature set for differentiating cancer from normal tissue.

Accuracy, sensitivity, specificity, and F-score as defined in [149, 186] were chosen as the metrics to evaluate the classification performance of the features. Accuracy represents the percentage of the correctly detected tumor and normal pixels relative to the total number of tumor and normal pixels in an image. Sensitivity represents the percentage of correctly detected tumor pixels relative to the total number of tumor pixels in an image. Specificity is the percentage of correctly identified non-tumor pixels relative to the total number of non-tumor pixels in an image. F-score is the harmonic mean of precision and sensitivity:

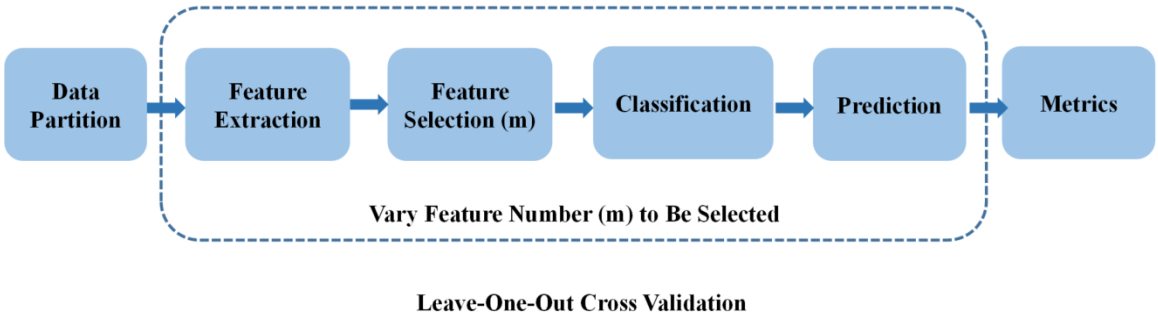


Figure 37: Flowchart for feature selection and classification.

5.4 Experimental Results

5.4.1 Results on Glare Detection and Removal

Figure 38 showed one example of glare detection results. The bright area in the standard deviation (std) image represented the location of glare pixels, which were characterized by large spectral variations and distributed mostly along the long tail of the histogram. The key issue was to identify an appropriate threshold on the histogram that could separate the glare pixels from non-glare pixels in the std images. We compared the traditional thresholding method such as the Ostu method [211] and entropy method [212] with the proposed loglogistic curve fitting method. The Ostu and entropy methods were most suitable for histograms with a bimodal shape, but the histogram here contained a very long tail which was not in the typical bimodal shape. The curve fitting methods produced a much lower threshold than the Ostu and entropy methods which enabled the detection of the relatively bright glare margins as well as the isolated glare pixels. The threshold was set to be the intensity value which yielded a certain percentage (ε) of the peak value in the fitted loglogistic distribution curve. The value of ε was experimentally set to 5% through trial and error, which was found sufficient to detect most of the glare pixels. After the glare masks were generated, glare pixels along all the spectral bands

were removed from the training and testing dataset since they contained no useful diagnostic information.

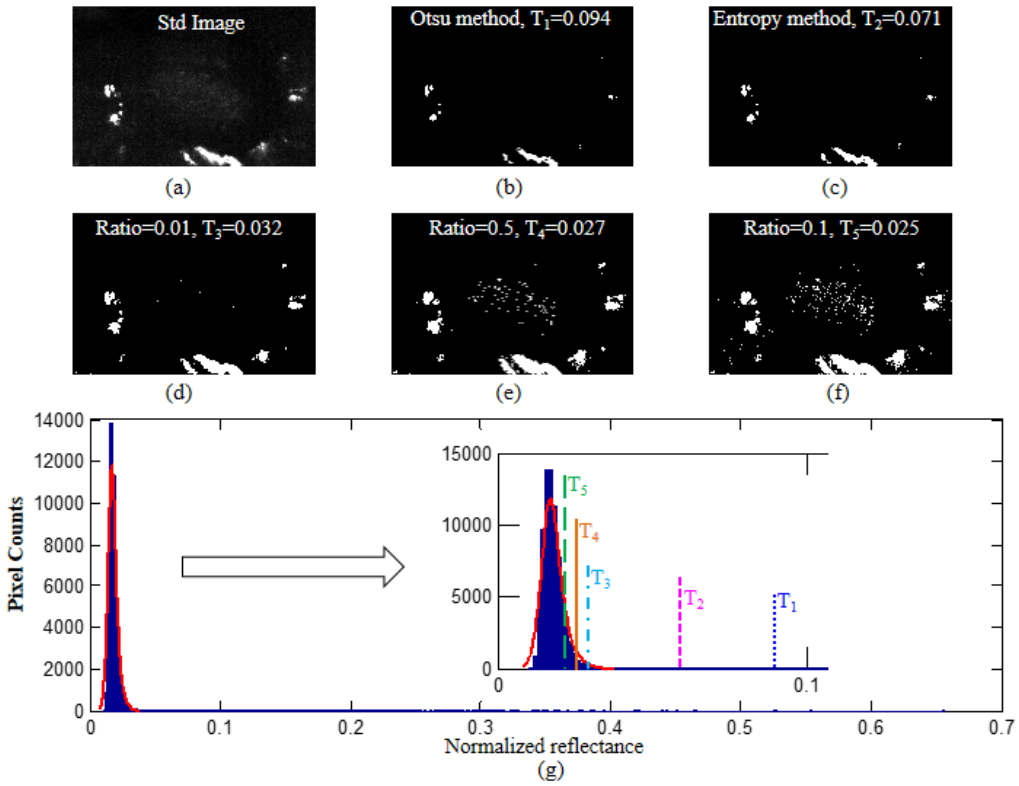


Figure 38: Glare detection results. (a) Standard deviation (std) image of the 1st order derivative for a hypercube. (b) Binary glare map generated by the classical Otsu method, (c) Binary glare map generated by the entropy method. (d)-(f)The glare maps generated by the proposed method with ratio 0.01, 0.05 and 0.1. Lower ratio led to under-detection of glare pixels, while higher ratio led to over-detection. 0.05 was the most suitable ratio. (g) Histogram of the std image pixels in blue with the loglogistic fitting curve in red. The five vertical lines represent the five thresholds generated by the five methods in (b)-(f).

5.4.2 Results on the Comparison between GFP and Non-GFP Images

Figure 39 and Figure 40 show the exposed tumors under the white excitation with and without GFP, respectively. GFP emission peaks under blue excitation occur at the wavelengths of 508 and 510 nm. It was found that under white excitation, the spectral

images at these two bands did not exhibit enhanced contrast between tumor and normal tissue compared to other spectral bands. This is consistent with the observation in our previous study [149]. Therefore, we did not remove GFP spectral bands at 508 and 510 nm in the pre-processing steps.

Both figures demonstrated that hyperspectral imaging can probe vessels at different depths below the visual surface of the tumor. Light at a shorter wavelength region of the visible range is more sensitive to superficial vascular information due to limited light penetration into the tissue. As the wavelength becomes longer, information from deeper tissue can be acquired. Hence, changing the illumination wavelength may enhance vascular contrast and allow visualization of angiogenesis at the tumor region.

The RGB color images show the highly vascularized tumors. Figure 39 and Figure 40 (c) illustrated the characteristic spectrum of hemoglobin at 542 nm and 577 nm. These characteristics may contribute to the distinction between cancerous and normal tissues by hyperspectral imaging.

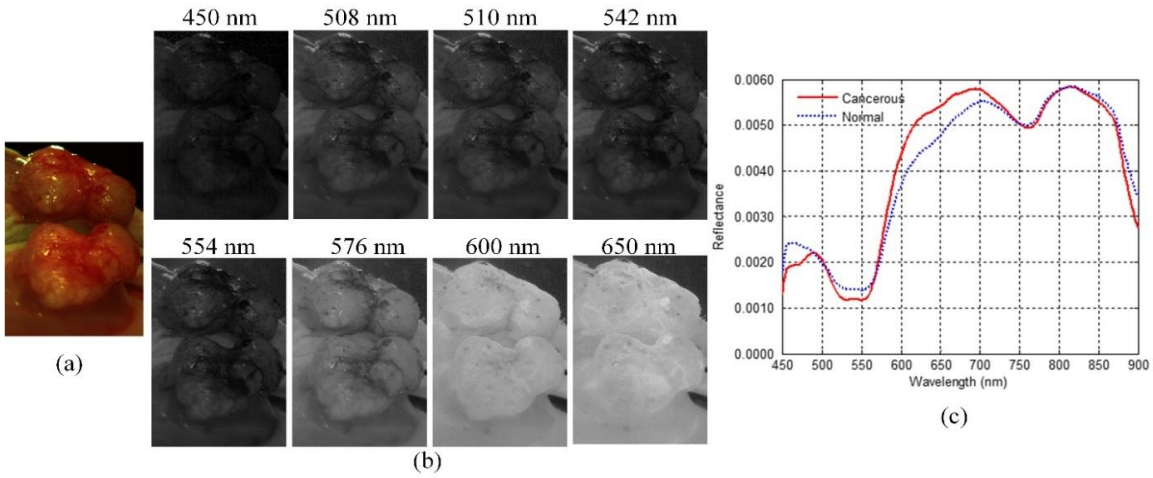


Figure 39: Visualization of a tumor with GFP. The upper part is the image of the tumor in a mirror. (a) RGB composite image of the hypercube, (b) Pre-processed spectral images at wavelength 450 nm, 508 nm, 510 nm, 542 nm, 554 nm, 576 nm, 600 nm, and 650 nm. (c) Spectral curve of cancerous and healthy tissue.

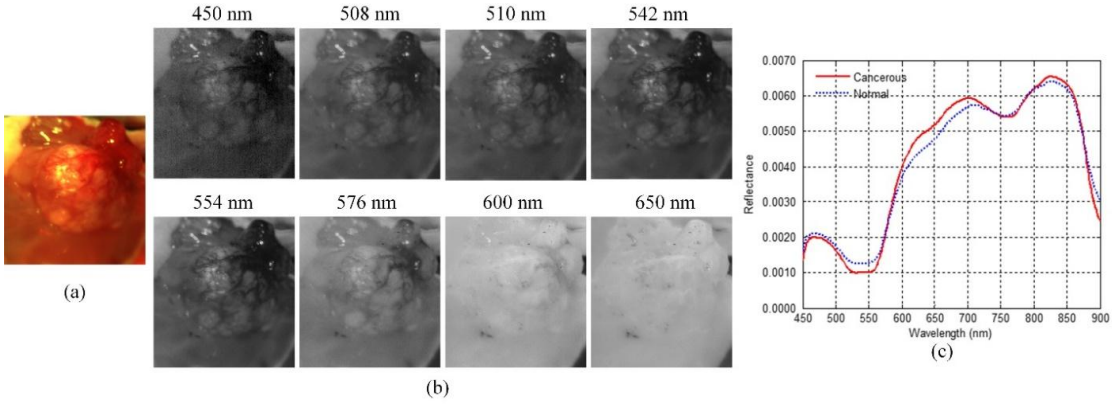


Figure 40: Visualization of a tumor without GFP. (a) RGB composite image of the hypercube, (b) Pre-processed spectral images at wavelength 450 nm, 508 nm, 510 nm, 542 nm, 554 nm, 576 nm, 600 nm, and 650 nm. (c) Spectral curve of cancerous and healthy tissue.

5.4.3 Results on Feature Extraction and Visualization

The most commonly utilized feature for cancer detection with hyperspectral imaging is the normalized reflectance spectra, which reflects the physiological and pathophysiological states of tissue at each pixel. However, this feature alone may not have enough discriminative information to minimize the classification error. It would be very interesting to explore the usefulness of other features besides reflectance. We derived a series of features based on the spectral curve of each pixel and boosted the original feature dimension from 226 to 904.

Figure 41 showed an RGB composite color image of an example hypercube. The tumor exhibited white necrotic appearance which was confirmed by the histological image in the interface between viable and necrotic tumor tissue. The reflectance spectrum of a tumor with necrosis differs from that of a tumor without necrosis. The image also shows that the pre-existing blood vessels in normal tissue reached the tumor region and grew new blood vessels into cancerous tissue. The highly vascularized tumor region grew well due to adequate oxygen and nutrient supply and the removal of waste through blood

vessel, while the non-vascularized tumor region became necrotic due to the lack of vessels. These observations were consistent with their spectra curves. The viable cancer tissue is heterogeneous with higher spectra variations compared to the necrotic tissue. The average reflectance spectrum of the viable tumor was lower than that of the normal tissue which reflected the higher amount of hemoglobin in the highly vascularized cancer tissue.

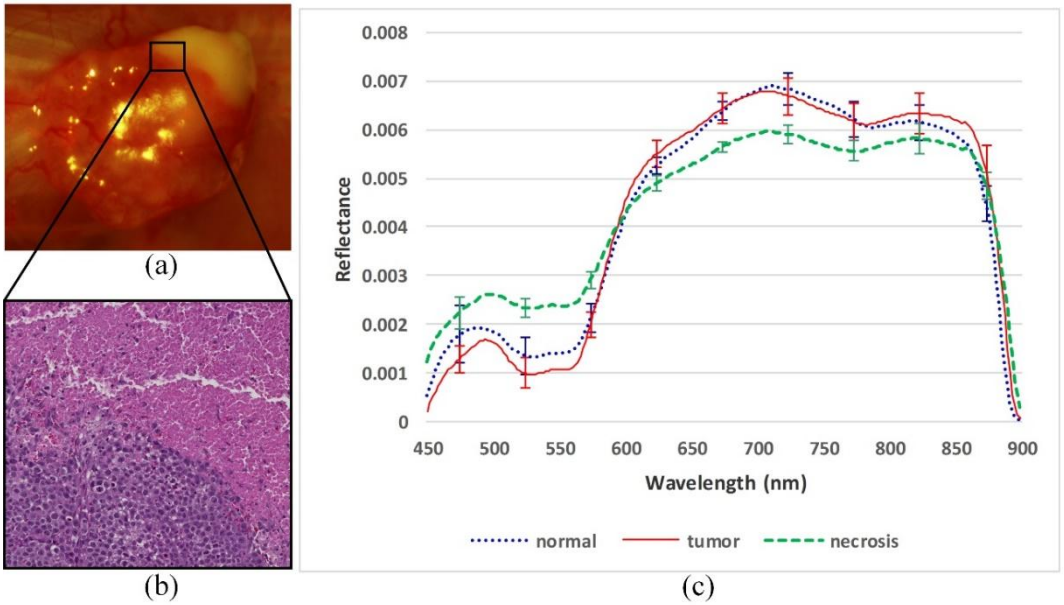


Figure 41: Reflectance spectral curve of a tumor with necrosis. (a) RGB composite image of a hypercube. The white region appears necrotic, and the other part of the tumor contains numerous vessels; (b) Histological image of the rectangular tissue region in (a). The upper region is the necrotic tissue without nuclei, and the lower part is the viable cancerous tissue; (c) The average reflectance spectra of the tumor, necrosis, and normal tissue with standard deviation. The red solid line represents the average spectra of cancerous tissue and the blue dotted line represents the average spectra of the normal tissue. The green dashed line represents the average spectra of the necrosis tissue. The error bars are the standard deviations at certain wavelength of the three curves.

Figure 42 showed the visualization of all the spectral features explored in this study, which include the RGB composite image of hypercube, and the mean, std, and sum

of the selected tumor and normal tissue ROI. The texture of the tissue could be clearly visualized from the std image. The mean and sum image only differ by a constant. Due to the large dynamic range of the spectra, along the wavelength, the average of the hypercube in the ROI did not reflect the correlation with the RGB image. All these features captured different aspects of the differences between tumor and normal tissue, which are visualized in these figures.

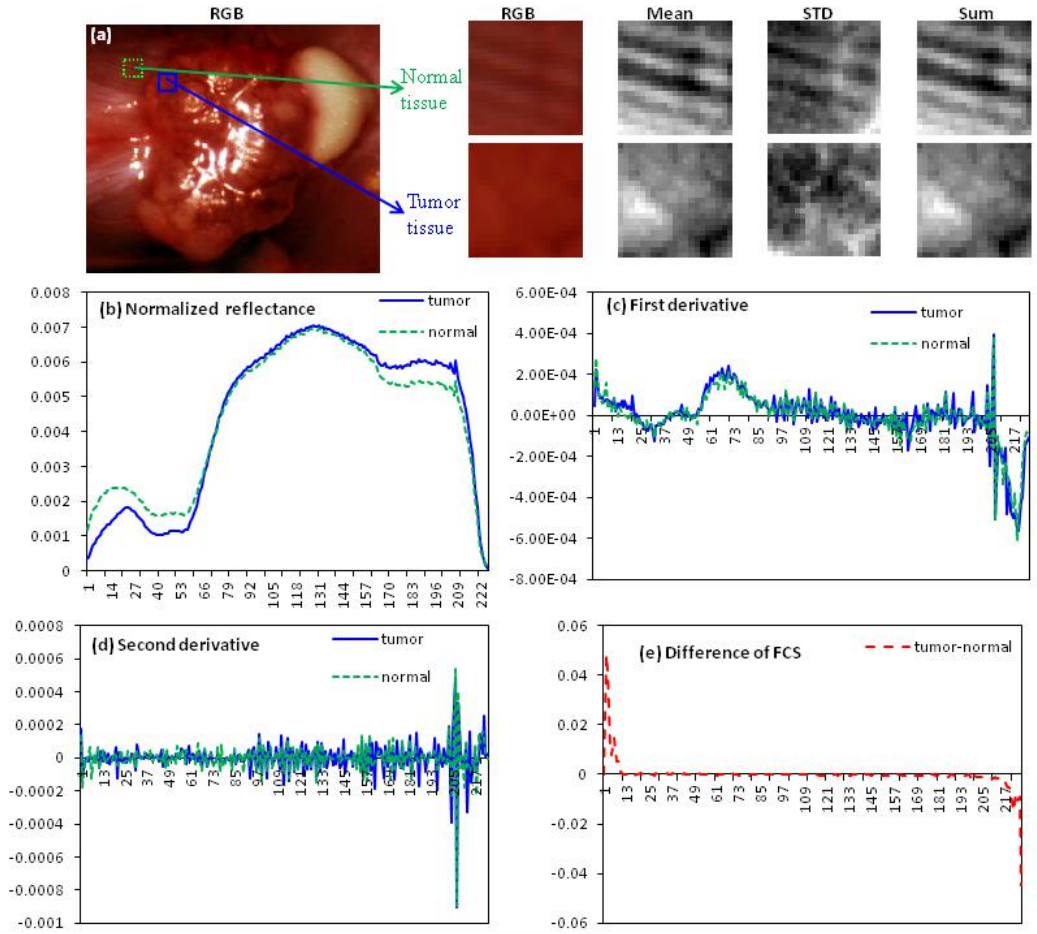


Figure 42: Feature extraction and visualization. (a) RGB composite image of a hypercube, along with the mean, std, and sum of the selected tumor and normal tissue ROI. (b-e) Average normalized reflectance curve, first derivative, second derivative and the difference of FCS between normal and tumor tissue in the selected ROI in (a).

5.4.4 Feature Selection and Classification Results

The objective of the feature selection is to identify the features, which is critical to minimize the classification error. The mRMR feature selection method selected a compact feature set with the maximal relevance to the target class and the minimal redundancy within the feature set. Figure 43 shows the mutual information of the extracted spectral features with the class labels (tumor or normal), which reflects the relevance of each feature with respect to the class labels. The highest mutual information was achieved by Fourier coefficients. Normalized reflectance above 850 nm showed higher relevance than other wavelengths. Derivatives, mean, std, and sum generally exhibited lower mutual information with class labels.

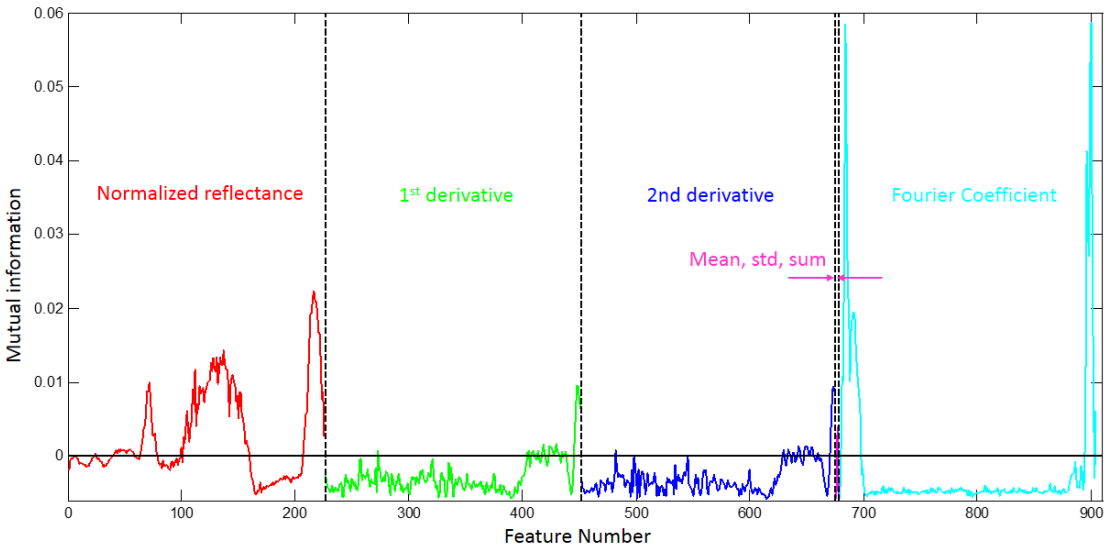


Figure 43: Mutual information between features and class labels. x axis represents the feature number and y axis represents the mutual information.

Figure 44 displays visualization of the mutual information among individual features which represents the redundancy within the 904 features. It can be seen clearly that the lower left square of size 226×226 shows relatively higher mutual information

than other regions which demonstrates that wavelength features are highly correlated with each other. Therefore, the normalized intensity values across the wavelength range contain complementary information as well as redundancies.

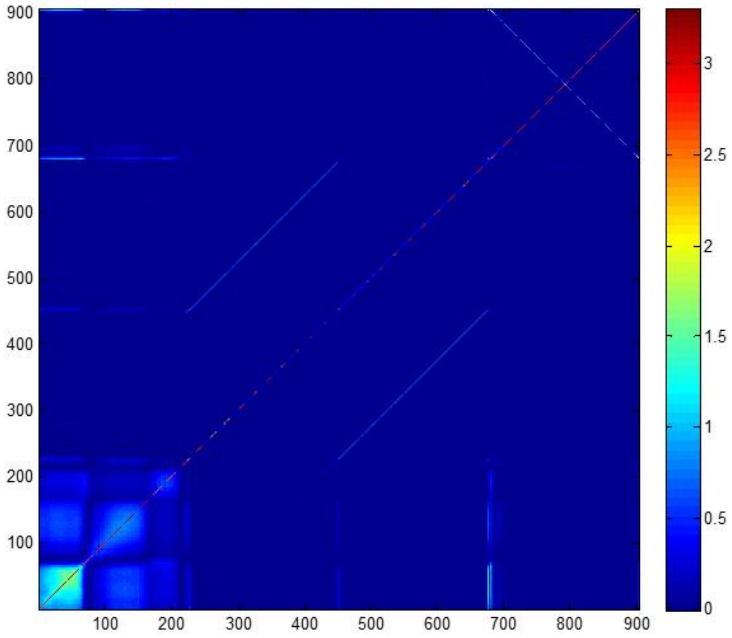


Figure 44: Mutual information between individual features. Color bar on the right shows the color map corresponds to the value of mutual information. Higher mutual information indicates more redundancy between features.

Figure 45 shows the evaluation of feature sets of different sizes by supervised classification. The metrics initially increased with the feature number, reached a maximum, and then decreased as the feature set went to its maximum size of 904. Classification with the full feature set of dimension 904 was not as good as the feature set of dimension 20. We found that a feature set of size 20 gave the best performance, with an average accuracy of 67.2%. In this experiment, classification was used only for evaluating and comparing the effectiveness of feature sets of different sizes. Since the

focus of this study was not on classification, we would try to improve classification in the future.

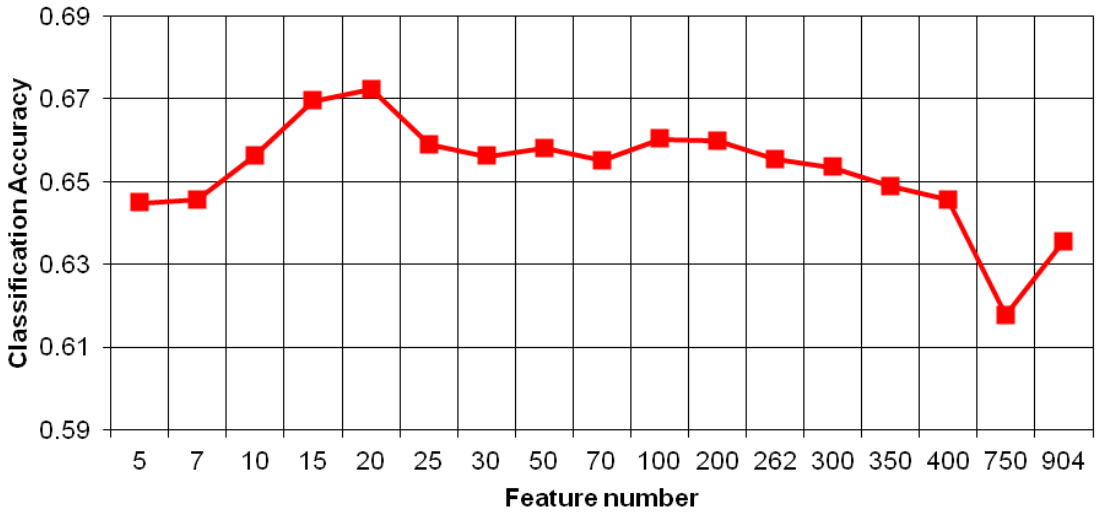


Figure 45: Feature selection and classification.

As shown in Figure 46, we have run mRMR with cross validation to select the optimal feature set $F = \{f_i\}_{i=1,2,\ldots m}$ of size m (m=1, 3, 5, …750, 904) from 904 features. We have generated a series of feature rankings: $\{f_1\}$, $\{f_1, f_2, , f_3\}$, …, $\{f_1, f_2, \ldots, f_{750}\}$, $\{f_1, f_2, \ldots, f_{904}\}$, where f_i represents the feature that has been ranked in the i^{th} position of each optimal feature set. It should be noted that f_i could be different on different cross validation folds or among optimal feature sets of different sizes. Therefore we can analyze the composition of the i^{th} ranked feature f_i and count the ranking frequency of individual features. Figure 46 summarized the ranking frequency of different feature types. Each bar represents the normalized frequency of different feature types being selected as the i^{th} ranked feature (i=1, 2, …, 20). The Fourier coefficient was the only selected top one ranking feature type. Normalized reflectance was the only feature type that ranked second, and the mean spectrum is the third-ranked feature type.

Spectral derivatives ranked the fourth and fifth even though their mutual information with corresponding labels was low. This could be explained by the fact that mRMR not only selected the features with high mutual information but also with low redundancy within the feature set.

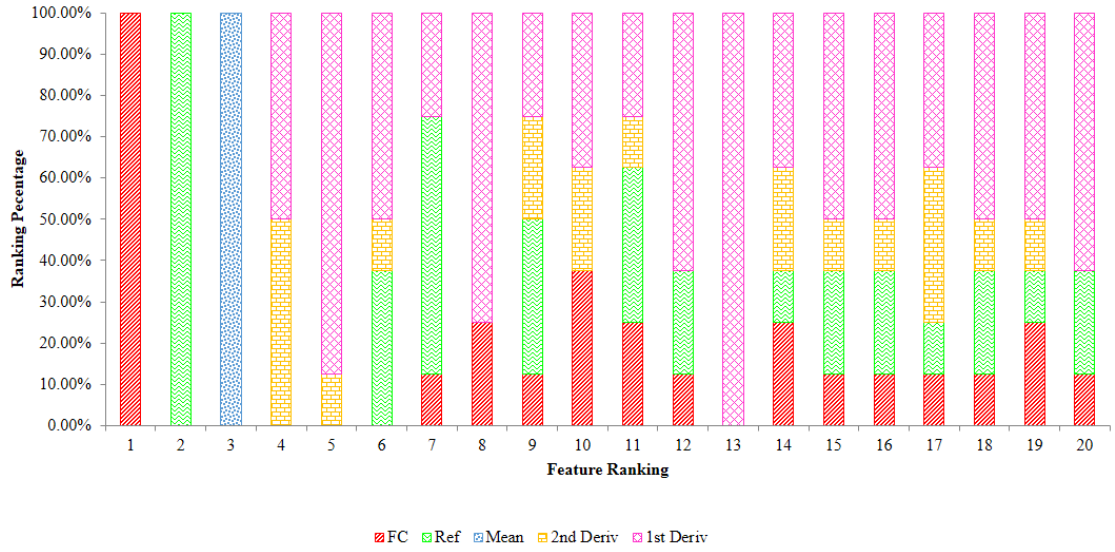


Figure 46: Feature ranking. X axis is the ranking from 1 to 20 and y axis is the percentage of the selection frequency for different features on each rank. Each bar represents the normalized frequency of different feature types being selected as the i^{th} ranked feature.

5.5 Discussion

Complete excision of tumors remains one of the key challenges in tumor surgery. Failure to remove all tumor cells increases the risk of tumor recurrence and the need for secondary surgery. Current intraoperative margin assessment is performed by visual inspection and palpation, followed by examination of suspicious areas using histopathological evaluation. Intraoperative cancer imaging and diagnosis is valuable to surgeons in the evaluation and excision of sequential layers of tissue. We developed pre-processing techniques and explored the relevance of different spectral features on tissue

labels and redundancy among individual features. Although the proposed method was tested in an animal model, the pre-processing and feature analysis techniques should also work well for the real clinical situation since the glares, motion artifacts, curvature, and high dimensionality problems may exists in any intraoperative hyperspectral images. We showed that the combination of Fourier coefficient, normalized reflectance, mean, and spectral derivatives could improve the distinction between cancerous and normal tissue. Our long-term goal is to provide visual guidance during surgery to aid tumor margin evaluation, thereby decreasing the amount of tumor left behind and increasing tumor-free survival. Although we used an open operative technique, this technology should be well suited to laparoscopic and robotic surgery by coupling hyperspectral imaging instruments with a laparoscope.

The key advantage of hyperspectral imaging for intraoperative tumor visualization is that it acquires images of a large area of tissue in a non-contact manner and generates objective tumor maps to enable surgeons to excise and evaluate sequential layers of tissue during surgery without the need for tissue sectioning and staining. The penetration depth of the hyperspectral imaging under 900 nm does not exceed a few mm, which is adequate for the sequential imaging and diagnosis during surgery. This technique could provide fast and objective feedback to the surgeon as to whether all tumor tissue has been excised or whether further tissue removal is required. Although this study used head and neck cancer as an example, HSI may be used to provide intraoperative diagnosis and ensure clear margin during tissue-conserving surgery for other cancer types.

One of the common problems during intraoperative imaging is that glare caused by specular reflection of the liquid on the tissue surface is always present in the surgical images, which deteriorates the quality of optical imaging and affects the quantitative analysis of the images. Cross-polarization has been widely used for the glare removal during the imaging stage by placing a linear polarizer in front of the light source and

another orthogonal polarizer in front of the detector. The orthogonal polarization has the effect of selectively detecting photons which have undergone multiple scattering events in the tissue and rejecting the photons from specular reflection from tissue surface. As there are no cross polarizers in our imaging instrument, we developed a simple but effective method to remove glare pixels and thus the diagnostic information contained in these pixels is also lost during the process. To the best of our knowledge, this is the first time that a fast and efficient glare detection method has been designed specifically for intraoperative hyperspectral images. The proposed glare removal method can also be applied to other optical images for eliminating glare regions.

In this study, GFP fluorescence images of tumors are used as the in vivo gold standard for tumor margin assessment. It was demonstrated that under white light illumination, spectral images at the GFP emission peaks do not show higher contrast than images at other wavelengths. Therefore, the GFP signal does not contribute to the differentiation of tumor and normal tissue by spectral-spatial classification. In the future, *ex vivo* histopathological images can be further registered to the *in vivo* images in order to provide a more detailed diagnostic map.

One of the challenges in applying medical hyperspectral imaging for surgical guidance is that the imaging and diagnosis procedures should be done in a clinically useful time frame [213]. The imaging instrument used in this study usually takes about 2 minutes to acquire 226 spectral images from 450 nm to 900 nm with 2 nm increment, and the image testing takes about 1 minute. The time cost for evaluating tissues during surgery is significantly reduced compared to traditional histopathology (twenty minutes to an hour). The diagnosis with hyperspectral imaging was performed by quantitative image analysis. A high feature dimension would lead to more computation costs and therefore it is essential to extract only the most relevant features for diagnosis. It should be noted that “the m best features are not the best m features”, which means that the combinations of individually good features do not necessarily lead to good classification

performance. As such, the best 20 features were not the 20 best features with highest mutual information. The best feature set should be relevant to the class labels as well as be complementary to each other. Although only spectral features were used in this study, spatial features such as texture descriptors could be incorporated to further improve the classification performance. In addition to mutual information, similarity metrics such as F-score and Pearson's correlation coefficient can also be used to characterize the relevance and redundancy. With the advancement of the hyperspectral imaging hardware and software, the imaging speed and image analysis time may be further reduced. Therefore, hyperspectral imaging is promising for intraoperative tumor margin delineation and visualization. In the future, we will proceed to evaluate the ability of hyperspectral imaging for residual tumor detection, which is also a key issue affecting the patient prognosis.

5.6 Conclusion

We developed a set of pre-processing techniques for improving the quality of intraoperative hyperspectral imaging data. We also evaluated our feature extraction and a set of selection methods for differentiating cancer from normal tissue using a head and neck cancer model. We further demonstrated the potential of hyperspectral imaging as a noninvasive tool for tumor visualization and classification during surgery in an animal study. Hyperspectral imaging and quantitative analysis methods could have a variety of applications in the future.

CHAPTER 6

QUANTITATIVE DETECTION OF HEAD AND NECK CANCER IN FRESH SURGICAL SPECIMEN OF HUMAN PATIENTS USING HYPERSPECTRAL IMAGING

Surgical resection is one of the main treatment options for head and neck cancers. The extent of cancer resection is commonly assessed during surgery by pathologic evaluation of the frozen sections at the resected specimen margins to verify whether cancer is present. However, this procedure is costly in terms of time and personnel and may suffer from sampling errors. In Chapter 5, we have demonstrated the feasibility of HSI and quantification framework for rapid cancer detection during an animal tumor surgery. The ultimate goal is to develop HSI as a diagnostic tool for tumor margin assessment during cancer surgery of human patients. To move forward, we conducted a proof-of-concept study to determine the feasibility of using HSI for identifying and differentiating tumor from normal tissue for a variety of anatomic sites of head and neck cancer patients. Tissue from 36 consented patients undergoing head and neck cancer surgery were scanned with HSI, Two quantitative analysis frameworks were proposed to differentiate cancer from normal tissue. The cancerous regions delineated by the automated classification methods were validated by pathological diagnosis. The experimental results suggested that HSI combined with machine learning techniques could be developed for the assessment of the tumor margins on surgical specimen.

6.1 Introduction

Surgical resection of cancer remains one of the major treatment options for this disease. The central objective of surgery is to maximize the removal of tumors while preserving vital anatomic structures and functions. Extensive resection with unnecessary

removal of normal tissue can leave patients with serious functional and aesthetic deficits, compromising their ability to perform vital daily functions such as chewing, swallowing, or speaking. However if the tumor margins are not accurately defined and the diseased tissue is not completely removed, cancer is likely to persist or recur. Positive surgical margins, defined as the presence of tumor cells at the resected specimen edge, are associated with a poor prognosis in terms of increased local recurrence and decreased overall survival [214]. Thus, the ability to define the margins of tumor with a high degree of accuracy is critical for maximizing the efficacy of surgical treatment and the patient's subsequent quality of life, which might produce significant cost savings.

Achieving resection margin adequacy is highly "operator dependent" with respect to surgeons and pathologists. The practice of quantifying resection margin distance on the pathology report has not been universally adapted [215]. Intraoperative frozen section (FS) assessment of surgical margins is widely used to assist in complete tumor extirpation in head and neck surgery. However, frozen section analysis is costly with as much as $3,123 on average spent per patient and an estimated cost-benefit ratio of 20:1 [216]. The overall accuracy of frozen section has been reported to be only 71.3% and 77.9% in the evaluation of close (< 5mm) or positive final margins in two comprehensive studies [216] [217]. Freezing artifacts such as distortion of tissue architecture, uneven sectioning, or poor staining increases the likelihood of interpretive errors of histologic diagnosis. Moreover, frozen section histopathology requires multiple well-trained professionals, increases the cost and the duration of procedures, extends anesthesia-related risks and provides diagnostic information at only a few discrete locations of the resection margins [217]. Therefore, intraoperative techniques that can be easily used in clinical practice to facilitate the fast and accurate assessment of surgical resection margins could improve the care delivered to patients with head and neck cancer.

HSI is a promising technique to assist surgeons in pinpointing the boundaries of cancers more precisely before and during surgery. Hyperspectral images, which contain

spectral signature at each image point, can be analyzed to identify various pathological conditions. With spatial information, the entire surface area of interest can be interrogated, potentially reducing the chance of sampling error and enabling a more thorough evaluation. One of the major advantages of this technique over other imaging modalities is that HSI has the ability to provide tissue information without using an exogenous molecular imaging agent by making use of endogenous contrast alone.

The objective of this study was to evaluate the potential of hyperspectral imaging to distinguish tumor from normal tissue on fresh surgical specimens of head and neck cancer patients. We proposed two quantitative analysis frameworks that utilize machine learning algorithms to differentiate cancer from normal tissue in hyperspectral images. The proposed detection methods were validated by both the pathological diagnosis and other optical imaging methods including autofluorescence imaging and fluorescence imaging of 2-NBDG and proflavine. To the best of our knowledge, this is the first time that HSI combined with machine learning was tested and validated in a wide variety of head and neck cancer tissue from human patients.

6.2 Materials and Methods

6.2.1 Hyperspectral Imaging Instrumentation

The HSI system used in this study has been described in Chapter 2.

6.2.2 Surgical Specimen Collection

Head and neck cancer patients scheduled to undergo cancer resection surgery at Emory Hospitals were preoperatively consented for our study. All tissues were collected under the Head and Neck Satellite Tissue Bank (HNSB, IRB00003208) protocol approved by the Emory University Institutional Review Board. The samples were de-identified and coded by a Clinical Research Coordinator before being released to our laboratory. During the resection surgery of head and neck cancer, fresh surgical

specimens were sent to the pathology room for assessment. Selected surgical margins were processed for frozen-section histopathologic evaluation, as routinely performed, independent of our research. When available, three tissue samples, clinically visible tumor, neighboring healthy tissue, and tumor-normal interface (tumor with adjacent normal tissue), were procured from the main specimen of each consented patient. Tissue regions with pathology ink were avoided. Specimen collection and imaging did not affect procedure time in the operating room or the content and verification of the final pathology report. Figure 47 illustrates an overview for experimental design of surgical specimen imaging.

6.2.3 Fresh Surgical Specimen Imaging

Fresh surgical specimens collected from the hospital were kept in cold PBS and immediately brought back to our imaging center. The resected specimen was washed carefully with PBS to eliminate the blood contamination on the tissue surface. A series of imaging steps were performed as follows.

Step 1: White and dark reference hypercubes were acquired before tissue imaging. White reference image cubes were acquired by placing a standard white reference board in the field of view. The dark reference cubes were acquired by keeping the camera shutter closed.

Step 2: The specimens were placed on a non-reflective black board on the imaging stage. Reflectance hyperspectral images of the specimens were obtained from 450-900 nm with 5nm intervals.

Step 3: Autofluorescence images were acquired with 455 nm excitation and 490 nm long-pass emission filters.

Step 4: Fluorescence images of 2-NBDG and proflavine were acquired. By topically applying 2-NBDG and proflavine to the surgical specimen, we may identify cancerous regions from fluorescence images and compare the fluorescence imaging with

label-free wide-field HSI. Detailed procedures for vital dye fluorescence imaging methods were previously described in Chapter 4.

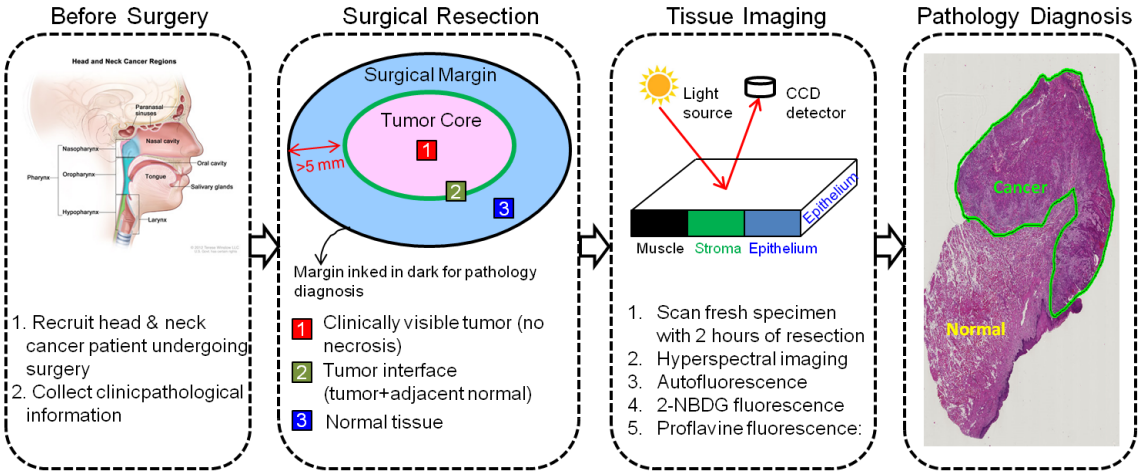


Figure 47: Overview of the experimental design for surgical specimen imaging

6.2.4 Histological Processing and Annotation

After all the imaging procedures were complete, we stained different sides of each tissue specimen with different color inks to preserve its orientation during histological processing. The inked specimen were then fixed in 10% buffered formalin overnight and sent to pathology department for standard histologic processing. Two to three serial sections of 5 µm-thicknesses from the imaging surface were cut and stained with H&E. These slides were digitally scanned for pathology diagnosis. A board-certified pathologist examined the histology slides and outlined the tumor border on the digitized slides as the gold standard to validate our quantification method.

6.2.5 Pre-processing of Hypercube

The method for data preprocessing consisted of reflectance calibration and glare removal:

Reflectance normalization: This step aims to remove the spectral nonuniformity of the illumination device and the influence of the dark current. The raw radiance data can be converted into normalized reflectance using Equation 1 as described in Chapter 2.

Glare Removal: This step removes glare pixels from the normalized hypercube. Glare regions were formed due to specular reflection from the moist tissue surface and did not contain any diagnostic information from under the tissue surface. The details of the method have been described in [218].

6.2.6 Feature Extraction

(1) Hyperspectral Imaging: To reduce computational time without reducing accuracy, spectral curves were averaged in non-overlapping blocks of N×N to yield a spectral signature per block. All of the spectral information available in the hyperspectral data was utilized. Blocks containing glare pixels were excluded from the classification process. Each block was assigned a label as cancerous or normal.

(2) Multispectral Imaging: For comparison, we evaluated the multispectral imaging (MSI) features utilizing the spectral peak bands of oxygenated hemoglobin (540 nm and 575 nm) and deoxygenated hemoglobin (555 nm), which have been shown to aid in the visualization of increased vasculature in the oral cavity during malignant transformation. The average multispectral intensities were used as features for MSI classification.

(3) Conventional RGB Imaging: We simulated the conventional RGB imaging by converting the illumination wavelength for each band in the range of 450-900 nm into the constituent RGB values as perceived by humans, then averaging the contribution to R, G, and B for each band. The average RGB intensity values of each block were used as features for RGB imaging classification.

(4) Autofluorescence Imaging: Average fluorescence intensity from 500 to 720 nm with 10 nm increments within each block was extracted as features for classification.

(5) Fluorescence Imaging: The mean fluorescence intensity of 2-NBDG and proflavine images from 500 to 720 nm with 10 nm increments within each block were extracted as features for classification.

6.2.7 Classification

Pathological Validation and ROI Selection

The gold standard for cancer detection is pathological diagnosis. Since the tumor margin was outlined in H&E stained histological images by a pathologist, we need to register the pathological image with the hypercube of the tissue specimen to delineate the tumor region in hyperspectral images which would be the gold standard for validating our quantification methods. To do this, we first synthesized an RGB color image from the hypercube and manually registered the histological image with the synthesized RGB image using affine registration. Due to tissue deformations (shrinkage, tearing, and distortion) during the histological preparation, it was very difficult to accurately align the outlined tumor boundary on the histological images with the actual tumor border on hyperspectral images. Therefore, only regions of interest (ROIs) that were histopathologically confirmed to be tumor or normal were selected for quantitative analysis.

Predictive Modeling

A total of 36 research subjects were included in this study (Table 10) for quantitative analysis. We tested and compared a variety of machine learning methods to analyze the spectra and determine classification criteria, allowing samples from all cancer tissue categories to be separated from samples corresponding to normal tissue. Classifiers including linear discriminant analysis (LDA) [64], quadratic discriminant analysis

(QDA), ensemble LDA [70], and linear support vector machines (SVM) [88] were used for supervised learning. These classifiers are described in detail in Chapter 2. All the data were processed and analyzed with software developed in-house that operates in a MATLAB environment (MATLAB 2015b, MathWorks, MA, USA). Based on our experimental design, two predictive analysis frameworks were developed and validated in this study as illustrated in Figure 48. Details of the methods are described below:

(1) Intra-patient Classification

As listed in Table 10, the tissues from 25 patients, including cancer from oral cavity (N = 12), larynx and pharynx (N = 3), maxillary sinus (N = 2), nasal cavity (N = 1), thyroid (N = 6), and parotid (N = 1), were analyzed with the intra-patient classification framework. The number in the brackets referred to the number of patients in this chapter. For each of these research subjects, samples were taken at the clinically visible tumor center without necrosis, surrounding normal tissue, and at the tumor - normal tissue interface. For each patient, intra-patient classification was conducted to detect and delineate the tumor regions in the tumor-normal interface sample. This classification method built training models with the spectral features extracted from the tumor and normal tissue samples, and evaluated the predictive performance of the model on the tumor-normal interface tissue sample of the same patient. The sensitivity and specificity of the classifier for each patient was calculated based on how many tumor pixels or blocks were correctly classified and how many normal pixels were correctly classified within the selected ROIs on the tumor-normal interface tissue. Utilizing spectra from patient-self for training avoided the influence of inter-patient heterogeneity.

(2) Inter-patient Classification:

To account for inter-patient heterogeneity, we proposed another framework called inter-patient classification. Tissue samples from 28 research subjects as listed in Table 10 were analyzed with this method. For each patient included for this analysis, at least two tissue samples (tumor and normal) need to be available. Leave-one-patient out cross-

validation was performed to account for the variances of classifiers. Seventeen patients with squamous cell carcinoma (SCC) or spindle cell squamous carcinoma (SCSC) or adenosquamous carcinoma (ASC), including cancers from oral cavity (N = 13), and larynx (N = 4), were grouped together as the first cohort for leave-one-out analysis. For these 17 cancer patients, during each run, 16 patients were randomly selected and all their tissue specimens were used as training dataset, while all the tissue specimens from the remaining patient were used as testing dataset. This process was repeated 17 times for the first cohort.

Similarly, all the 11 thyroid cancer patients, including 9 papillary thyroid carcinomas (PTC), 1 follicular thyroid carcinoma (FTC), and 1 medullary thyroid carcinoma (MTC), were considered as the second cohort for this analysis. For the 11 thyroid cancer patients, during each run, 10 patients were randomly chosen for training and the remaining patient for testing. This procedure was repeated 11 times for this cohort.

For each patient, we calculated how many normal pixels/blocks were correctly classified for a normal specimen, how many tumor pixels/blocks were correctly classified for a tumor specimen, as well as the sensitivity and specificity on a tumor-normal interface specimen.

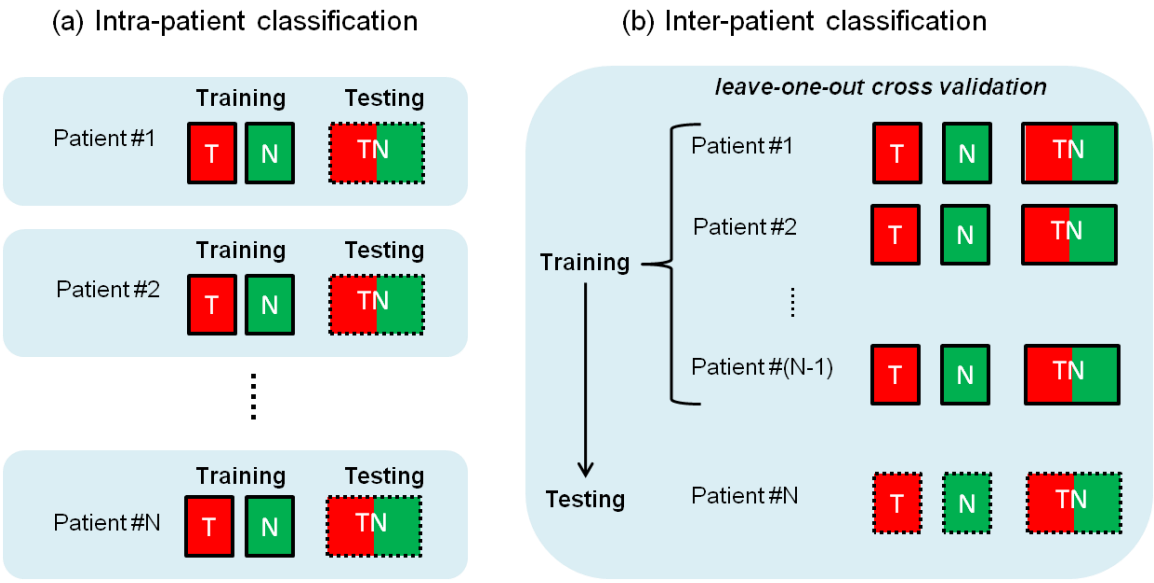


Figure 48: Overview of the two proposed quantitative analysis frameworks: (a) Intra-patient classification; (b) Inter-patient classification.

Table 10: Summary of surgical specimen and patient information for quantitative analysis (Gender*: F-Female, M-Male; Race*: A-Asian; AA-African American; H-Hispanic; I-Indian; W-White; Histologic type*: SCC-Squamous Cell Carcinoma; SCSC-Spindle Cell Squamous Carcinoma; ASC-Adenosquamous Carcinoma; Recur*****: Recurrent cancer; Intra*****: Patient selected for intra-patient classification; Inter*****: Patient selected for inter-patient classification)

Primary Tumor Site	**Specimen Origin**	**ID**	**Age**	**Gender ***	**Race ***	**Histologic type***	**Recur ***	**Intra ***	**Inter ***
Oral cavity	Tongue	1	55	F	W	SCC		√	√
	Tongue	2	43	M	W	SCC		√	√
	Tongue	3	67	F	W	SCC		√	√
	Tongue	4	75	F	W	SCC		√	√
	Tongue	5	60	M	I	SCC		√	
	Tongue	6	86	F	W	SCC	√	√	√
	Tongue	7	58	F	W	SCC	√		√
	FOM	8	50	M	W	SCC			√
	FOM	9	51	F	W	SCC		√	√
	FOM	10	57	M	W	SCC		√	√
	FOM	11	62	M	W	SCSC		√	
	Soft Palate	12	61	F	AA	SCC	√	√	
	Mandibule	13	53	M	W	SCC		√	√
	Mandibule	14	44	M	W	SCC			√
	Gingiva	15	76	M	I	SCC		√	√
	Alveolar Ridge	16	81	F	AA	SCC			√
Larynx	Supraglottis	17	44	M	AA	SCC		√	√
	Supraglottis	18	54	F	W	SCC			√
	Glottis	19	57	M	AA	SCC			√
	Glottis	20	69	M	AA	SCC	√	√	
	Hypopharynx	21	66	M	W	SCC		√	√
Paranasal &Nasal	Maxillary Sinus	22	73	F	W	SCC		√	
	Maxillary Sinus	23	65	M	W	SCC		√	
	Nose	24	74	M	W	ASC		√	

Table 10: Summary of surgical specimen and patient information for quantitative analysis (Continued)

Primary Tumor Site	Specimen Origin	ID	Age	Gender *	Race*	Histologic type*	Recur *	Intra*	Inter*
Gland	Thyroid	25	69	M	AA	PTC		√	√
	Thyroid	26	59	M	A	PTC		√	√
	Thyroid	27	24	F	I	PTC		√	√
	Thyroid	28	37	M	I	PTC		√	√
	Thyroid	29	22	M	H	PTC			√
	Thyroid	30	30	F	AA	PTC		√	√
	Thyroid	31	18	M	W	FTC			√
	Thyroid	32	37	F	AA	PTC			√
	Thyroid	33	35	M	W	PTC			√
	Thyroid	34	71	M	W	PTC		√	√
	Thyroid	35	35	F	W	MTC			√
	Parotid	36	39	M	AA	PA		√	

6.2.8 Performance Metric

We evaluated the performance of classifiers with receiver operating characteristic (ROC) curves, the areas under the curve (AUC), accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). Classification accuracy, sensitivity, specificity, PPV, and NPV were determined using the following equations (TN: true negative, TP: true positive, FP: false positive, FN: false negative):

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN};$$

$$\text{Sensitivity} = \frac{TP}{TP + FN}; \text{Specificity} = \frac{TN}{TN + FP};$$

$$\text{PPV} = \frac{TP}{TP + FP}; \text{NPV} = \frac{TN}{TN + FN}$$

6.3 Results

6.3.1 Spectral Visualization

Figure 49 shows spectral curves of tumor and normal tissue samples from a variety of head and neck cancer sites of human patients. We found that the measured average spectra for all cancerous tissue showed higher reflectance intensity than normal tissue from 450 to 900 nm in the case of oral cavity, thyroid, and larynx surgical specimen. Note that the above observations are valid on average for most cases. However, the spectra showed large within-class variations as well, which may primarily be attributed to the heterogeneity of head and neck tissue. The characteristic dip of hemoglobin can be observed in reflectance spectra of all cancer sites. In general, the mean reflectance intensity of tumor tissue is higher than that of normal tissue in most of these anatomical sites. In the case of parotid, adenoma tissue had lower reflectance than normal tissue in the middle wavelength region.

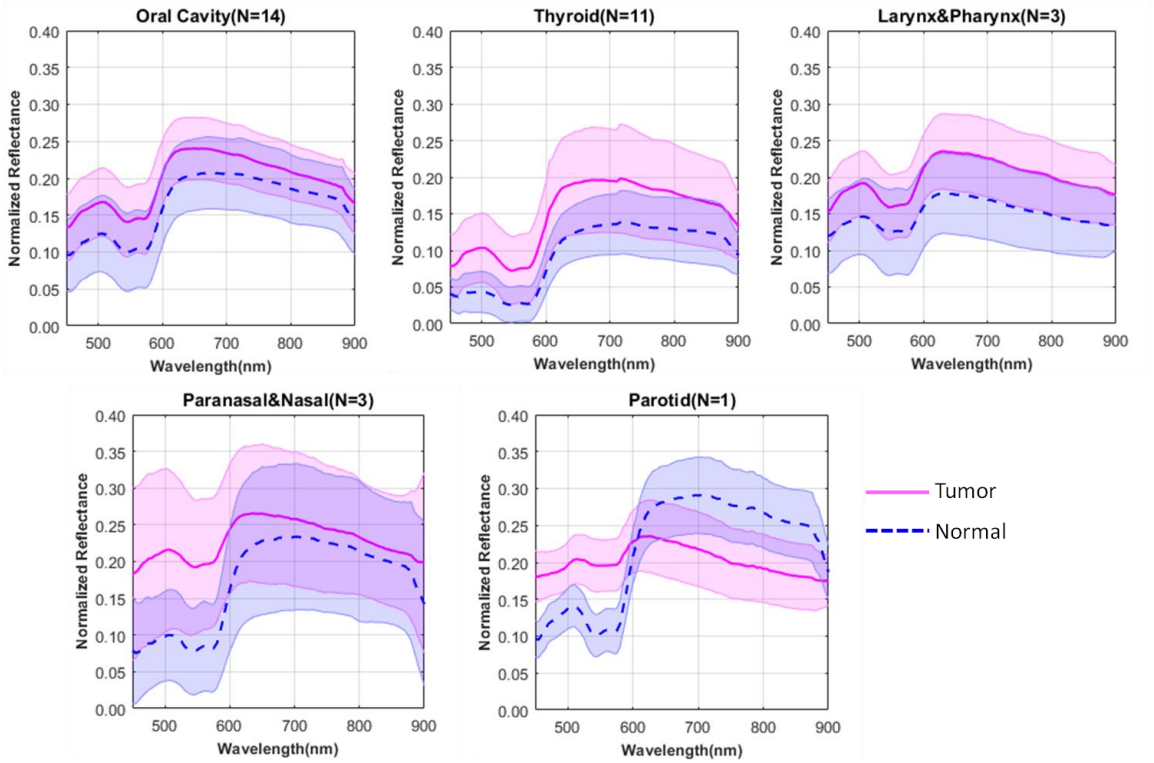


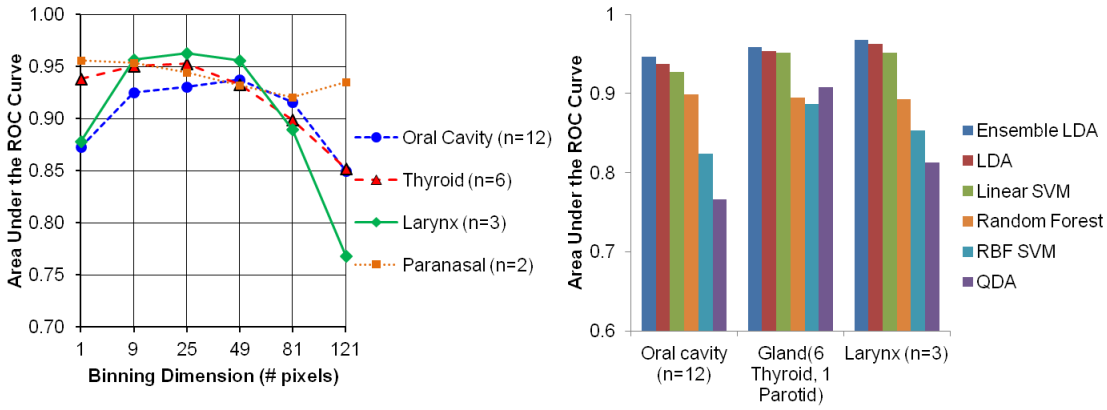
Figure 49: Average spectral curve of tumor and normal tissue samples from various head and neck cancer sites, including oral cavity, thyroid, larynx, pharynx, parotid, paranasal sinus, and nasal cavity of human patients. The magenta solid line and blue dashed line represent the mean spectra of cancer and normal tissue, and the shaded area centered on the two lines represent the standard deviation.

6.3.2 Intra-patient Classification

Tissue from different anatomic sites may have different structure and composition, so the optimal block size for best characterizing the tissue could be different. As shown in Figure 50 (a), the diagnostic performances of spectral features extracted from block size of 1×1, 3×3, 5×5, 7×7,9×9,and 11×11 were compared for specimens from oral cavity, thyroid, larynx, and paranasal sinus. The optimal block size for oral cavity is 7×7, for thyroid and larynx is 5×5, for paranasal sinus is 1×1. Among LDA, QDA, ensemble LDA, and linear SVM, ensemble LDA was chosen for classification due to its superior overall performance (Figure 50 (b)). As reported in Table 11, using reflectance spectra from HSI, we were able to distinguish between tumor and

normal tissue with an average accuracy, sensitivity, and specificity of 89%, 90%, and 90% from oral cavity, 91%, 91%, and 93% from gland (thyroid and parotid), 94%, 95%, and 90% from larynx and pharynx, and 90%, 90% and 90% from paranasal and nasal tissue, respectively. Overall, HSI outperformed autofluorescence imaging, as well as 2-NBDG and proflavine fluorescence imaging in all these cancer sites.

Next, we grouped the wavelengths from 450 to 900 nm into different sub-regions, including 450-600 nm, 605-850 nm, 855-900 nm and compared their classification performance. As shown in Table 12, we found that the whole wavelength region including both visible and near-infrared light performed the best for differentiating cancer from normal tissue in oral cavity cancers, while the visible wavelength region from 450 to 600 nm yielded the highest average classification accuracy for thyroid cancers.



(a) (b)

Figure 50: Comparison of diagnostic performance of different block sizes (a) and different classifiers for the distinction of tumor from normal tissue in multiple anatomic sites of head and neck cancer patients.

Table 11: Intra-patient classification performance of HSI, autofluorescence imaging, 2-NBDG fluorescence imaging, and proflavine fluorescence imaging.

Cancer Site	Imaging Method	AUC	Accuracy	Sensitivity	Specificity
Oral cavity (n=12)	**HSI**	**0.95±0.05**	**89%±8%**	**90%±7%**	**90%±8%**
	Autofluorescence	0.82±0.20	81%±16%	80%±16%	80%±19%
	2-NBDG	0.83±0.14	79%±14%	79%±15%	79%±14%
	Proflavine	0.68±0.18	66%±14%	64%±16%	69%±15%
Gland(6 Thyroid, 1 Parotid)	**HSI**	**0.96±0.04**	**91%±7%**	**91%±8%**	**93%±6%**
	Autofluorescence	0.72±0.31	67%±33%	80%±20%	71%±35%
	2-NBDG	0.84±0.18	80%±17%	78%±19%	82%±16%
	Proflavine	0.80±0.23	78%±21%	73%±26%	82%±15%
Larynx & Pharynx (n=3)	**HSI**	**0.97±0.03**	**94%±4%**	**95%±4%**	**90%±7%**
	Autofluorescence	0.74±0.26	74%±21%	74%±28%	79%±22%
	2-NBDG	0.97±0.04	92%±7%	92%±8%	92%±5%
	Proflavine	0.79±0.23	77%±20%	77%±20%	74%±23%
Paranasal & Nasal (n=3)	**HSI**	**0.96±0.02**	**90%±5%**	**90%±5%**	**90%±4%**
	Autofluorescence	0.81±0.11	76%±10%	79%±8%	72%±14%
	2-NBDG	0.76±0.18	68%±19%	64%±20%	79%±7%
	Proflavine	0.80±0.08	73%±6%	69%±11%	79%±2%

Table 12: Classification performance of wavelength sub-regions of HSI for the detection of head and neck cancer with intra-patient classification.

Primary Site	Wavelength Range (nm)	AUC	Accuracy	Sensitivity	Specificity
Oral cavity (n=12)	450-600	0.87±0.11	82%±11%	81%±13%	82%±10%
	605-850	0.89±0.12	84%±12%	83%±12%	85%±12%
	855-900	0.76±0.14	72%±11%	73%±11%	71%±12%
	450-900	**0.95±0.05**	**89%±8%**	**90%±7%**	**90%±8%**
Thyroid (n=6)	450-600	0.97±0.06	93%±8%	93%±9%	95%±6%
	605-850	0.95±0.06	91%±9%	90%±10%	92%±7%
	855-900	0.93±0.09	89%±10%	88%±12%	90%±8%
	450-900	**0.96±0.05**	**92%±7%**	**92%±9%**	**94%±6%**

To further demonstrate the potential of HSI for real-time guidance of surgical resection, we generated a cancer probability map with the tumor border outlined in green

as shown in Figure 51 and Figure 52. Figure 51 shows an example of tongue cancer detection and Figure 52 displays an example of thyroid cancer detection with intra-patient classification. In both cases, a predictive model was trained on the cancer and normal tongue tissue as and tested on the cancer interface tissue of the same patient. This color-coded map could assist surgeon to better visualize and assess tumor borders during surgery.

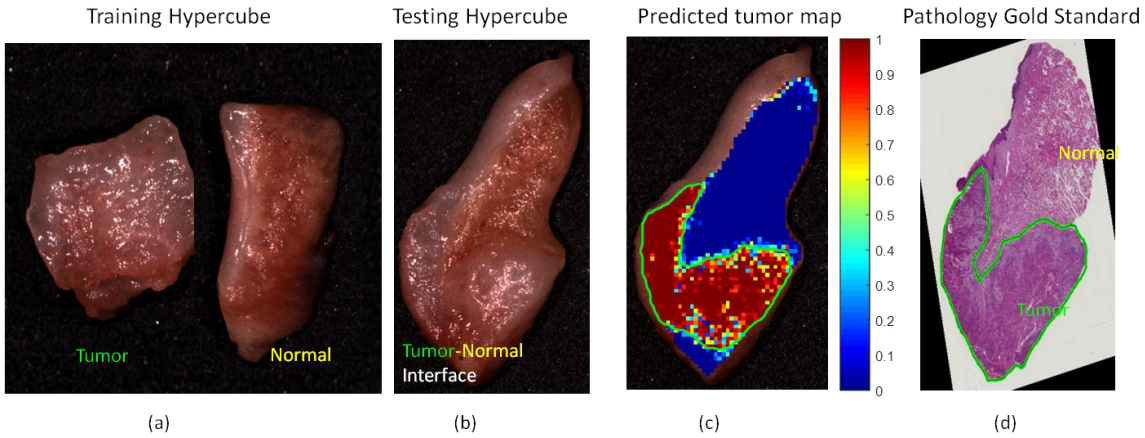


Figure 51: A representative tongue cancer detection result. (a) Training hypercube with a tumor specimen and a normal specimen. (b) Testing hypercube with tumor and normal interface tissue. (c) Cancer probability map generated by ensemble LDA classifier. Green line is the tumor border generated by thresholding on the probability map. The color bar shows the likelihood of being cancerous tissue. (d) Pathology gold standard with tumor region outlined within the green region by an experience pathologist.

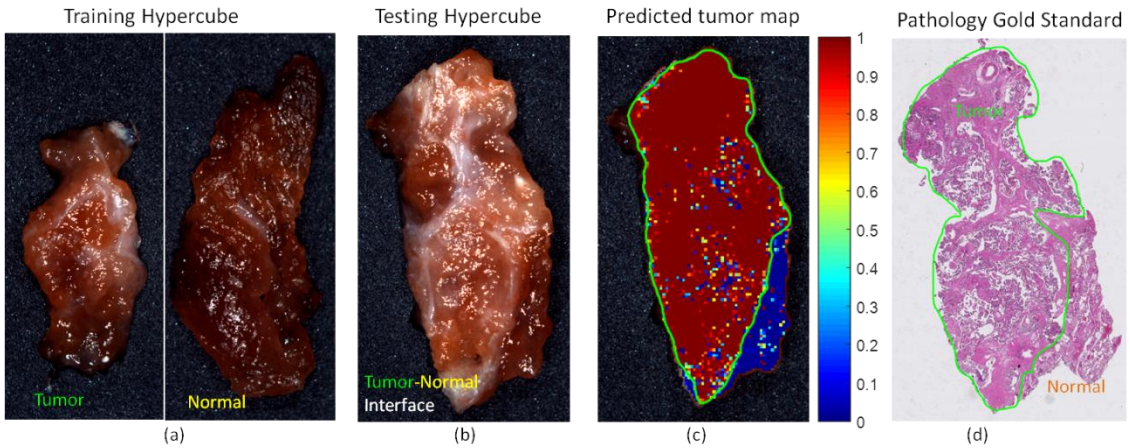


Figure 52: A representative thyroid cancer detection result. (a) Training hypercube with a tumor specimen and a normal specimen. (b) Testing hypercube with tumor and normal interface tissue. (c) Cancer probability map generated by ensemble LDA classifier. Green line is the tumor border generated by thresholding on the probability map. The color bar shows the likelihood of being cancerous tissue. (d) Pathology gold standard with tumor region outlined within the green region by an experience pathologist.

6.3.3 Inter-patient Classification

Figure 53 plotted the ROC curves of individual cancer patients with intra-patient classification (a) and inter-patient classification (b). With reflectance spectra from HSI, we obtained an average AUC of 0.88, and 0.91 for the two cohorts of patients described in section 6.2.7. The diagnostic performance of HSI for cancer detection was slightly lower than the results of intra-patient classification, probably due to the inter-patient heterogeneity.

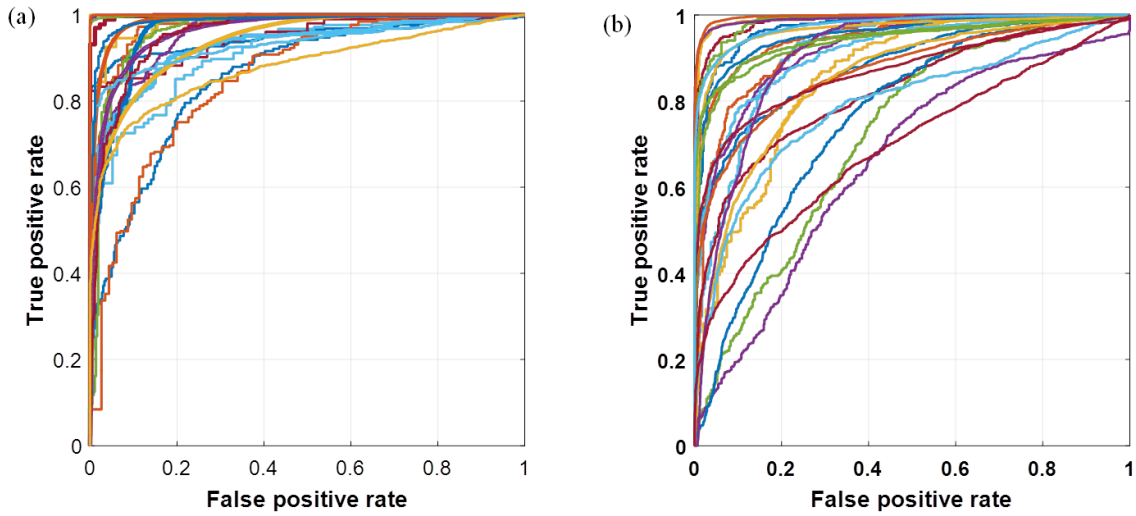


Figure 53: ROC curves of intra-patient classification (a) and inter-patient classification (b) with HSI for individual patients.

Similar to the intra-patient classification, we grouped the wavelength range of 450 to 900 nm into three sub-regions of 450-600 nm, 605-850 nm, and 855-900 nm, and compared their diagnostic performance with the inter-patient classification method (Table 13). We found that the visible wavelength region from 450 to 600 nm still provided significant diagnostic information, but the full wavelength region from 450 to 900 nm was able to distinguish cancer from normal tissue from oral cavity with higher sensitivity and specificity than the visible wavelength region alone. Moreover, we found that HSI was superior to MSI and conventional RGB imaging in oral cancer detection, as shown in Table 14.

Table 13: Classification performance of wavelength sub-regions of HSI for detection of head and neck cancer with inter-patient classification.

Primary Site	Wavelength Range (nm)	AUC	Accuracy	Sensitivity	Specificity
Oral cavity (n=13)	450-600	0.88±0.12	82%±11%	82%±11%	81%±12%
	605-850	0.82±0.11	77%±11%	78%±10%	76%±12%
	855-900	0.71±0.11	67%±9%	67%±9%	66%±11%
	450-900	**0.91±0.10**	**85%±10%**	**85%±9%**	**84%±12%**
Thyroid (n=11)	450-600	0.92±0.07	87%±9%	86%±9%	88%±9%
	605-850	0.94±0.05	88%±6%	88%±6%	88%±6%
	855-900	0.84±0.11	78%±11%	77%±10%	79%±10%
	450-900	**0.91±0.09**	**86%±10%**	**85%±11%**	**87%±9%**

Table 14: Comparison of HSI, MSI, and conventional RGB imaging for inter-patient classification.

Primary Site	Imaging Type	AUC	Accuracy	Sensitivity	Specificity
Oral cavity (n=13)	**HSI**	**0.91±0.10**	**85%±10%**	**85%±9%**	**84%±12%**
	MSI	0.77±0.19	73%±15%	73%±14%	73%±19%
	RGB	0.76±0.19	73%±14%	73%±14%	72%±16%
Thyroid (n=11)	**HSI**	**0.91±0.09**	**86%±10%**	**85%±11%**	**87%±9%**
	MSI	0.94±0.06	88%±8%	88%±8%	89%±7%
	RGB	0.95±0.04	89%±5%	90%±5%	89%±6%

Next, we looked at the cancer prediction accuracy for individual surgical specimens, including purely normal, purely cancer, and cancer-normal interface. As shown in Table 15, for patient within the first cohort (oral cavity, and larynx and pharynx), HSI achieved an accuracy of 82% in classifying cancerous regions in cancer tissue and an accuracy of 87% in classifying normal regions in normal tissue and. In addition, HSI was able to distinguish cancer from normal tissue with an average accuracy, sensitivity, and specificity of 83%, 68%, and 77% in tumor-normal interface tissue. For patients with thyroid carcinoma, HSI was able to accurately detect 86% of the

cancerous tissue from cancer specimen, and87% of the normal tissue from normal specimen as shown in Table 16. For cancer interface tissue, HSI achieved an accuracy, sensitivity, and specificity of 85%, 86%, and 87% respectively for the detection of thyroid carcinoma. Specimen from only one patient of the thyroid cohort was medullary thyroid carcinoma, so the classification results for this patient (P35) was relatively poor.

Table 15: Classification accuracy of HSI with inter-patient classification method for purely normal, purely cancer, and cancer-normal interface tissue specimens from the first patient cohort (Empty space in the table means no such sample).

Cancer Site	**PID**	**Tumor Tissue**	**Normal Tissue**	**Tumor-Normal Interface**		
		Accuracy	**Accuracy**	**Sensitivity**	**Specificity**	**Accuracy**
Oral Cavity	P1	99%	80%	48%	99%	68%
	P2	88%	89%	92%	90%	91%
	P3	99%	86%	92%	99%	97%
	P4	83%	91%	90%	58%	69%
		-	96%	-	-	-
	P5	-	59%	66%	73%	69%
		-	-	89%	39%	71%
	P6	90%	80%	-	-	-
	P7	-	97%	95%	66%	88%
	P8	-	-	84%	93%	91%
		-	-	99%	85%	91%
	P9	-	80%	74%	76%	75%
		-	94%	93%	88%	91%
	P10	77%	97%	87%	54%	77%
	P11	67%	84%	65%	21%	46%
	P12	91%	99%	98%	75%	90%
	P13	80%	89%	-	-	-
		-	87%	-	-	-
Larynx and Pharynx	P14	61%	88%	97%	41%	83%
	P15	-	91%	80%	51%	68%
		-	-	52%	43%	44%
	P16	71%	-	98%	81%	87%
	P17	78%	80%	87%	56%	77%
Average		**82%**	**87%**	**83%**	**68%**	**77%**

Table 16: Classification accuracy of HSI with inter-patient classification method for purely normal, purely cancer, and cancer-normal interface tissue specimens from the second patient cohort (thyroid carcinoma).

Cancer Type	PID	Cancer Tissue	Normal Tissue	Tumor-Normal Interface		
		Accuracy	Accuracy	Sensitivity	Specificity	Accuracy
PTC	P25	-	-	99%	94%	97%
		-	-	94%	100%	96%
	P26	93%	98%	76%	85%	80%
	P26	-	74%	71%	51%	59%
		-	-	68%	98%	96%
	P28	59%	86%	90%	87%	90%
	P29	88%	92%	-	-	-
	P30	97%	95%	98%	99%	99%
	P31	74%	93%	-	-	-
		97%	-	-	-	-
		97%	-	-	-	-
	P32	86%	78%	-	-	-
		-	93%	-	-	-
	P33	89%	89%	-	-	-
FTC	P34	99%	97%	76%	76%	76%
MTC	P35	64%	65%	90%	87%	90%
Average		**86%**	**87%**	**85%**	**86%**	**87%**

Figure 54 shows an example of thyroid cancer detection from hyperspectral images with inter-patient classification. We noticed a much lower reflectance spectrum from normal thyroid tissue than thyroid cancer tissue. We also found that cancer and normal tissue were accurately detected in all three tissue samples (cancer, normal, and cancer interface).

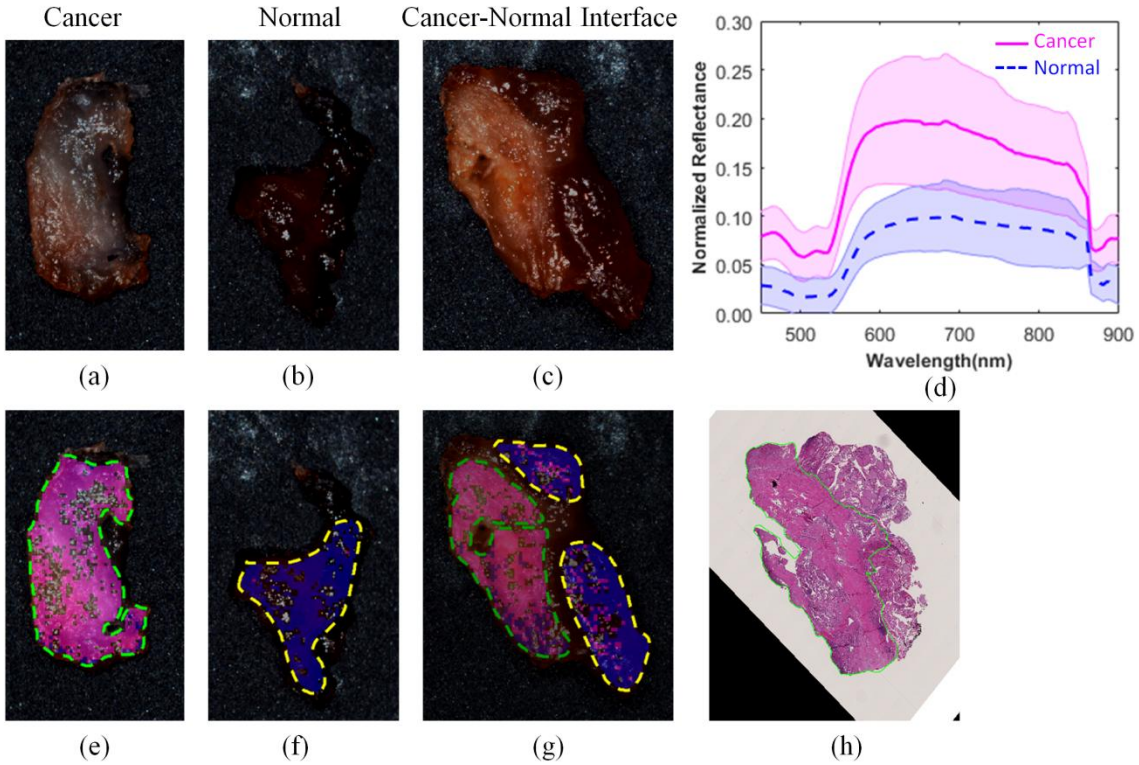


Figure 54: Example thyroid cancer detection with inter-patient classification. (a-c) are the synthesized RGB images from corresponding hypercube of cancer, normal, and cancer-normal interface surgical tissue procured from a patient with papillary thyroid carcinoma undergoing total thyroidectomy. (d) is the average spectral curve of cancer and normal tissue from the same patient, with shaded region representing the standard deviation. (e-g) are the prediction results with inter-patient classification method. Green and yellow curves outline the regions diagnosed as cancer and normal respectively. Magenta and blue colors denote predicted cancer and normal tissue regions. Glare pixels excluded from classification are not labeled. (h) is the registered pathology image with tumor region outlined inside the green curve by an experienced pathologist.

6.4 Discussion

We investigated the diagnostic potential of hyperspectral imaging for head and neck cancer detection in fresh surgical specimen from a variety of anatomic sites, including oral cavity, thyroid, larynx, parotid, paranasal sinus, and nasal cavity. We show here that HSI is capable of accurate, sensitive, and specific tissue classification of head and neck cancers. More than 1.7 million reflectance spectral signatures were obtained by HSI and analyzed with machine learning methods. The reflectance spectra capture the

alteration of absorption and scattering properties of tissue associated with malignant transformations. Molecular fingerprinting based on inverse modeling of reflectance spectra obtained by HSI may shed new light on our understanding of human head and neck cancer biology. Here, however, we are reporting on automated tissue classification methods that are using the full spectrum of 450 to 900 nm containing all of the molecular information to provide a diagnostic measure. This study demonstrates the feasibility of the application of label-free HSI as a promising tool for surgical margin assessment. Although frozen section diagnosis is commonly used to guide surgical resection during surgery, it only samples a small portion of the resection margin (approximately 0.1% to 1%), which may lead to underestimation and does not guarantee margin-negative resection. Moreover, this procedure is time-consuming and labor-intensive. HSI is a wide-field modality that is capable of covering the entire margin. It is able to sense tumors in varying depth with VIS and NIR light illumination (at clinically relevant sensing depths). The scanning time is on the order of minutes. It is an objective, fast, and cost-effective tool that may provide real-time assessment of complete resection margins. This technique is able to quantify not only surface mucosa but also different faces of the resection margins.

In this study, label-free HSI was shown to be superior to autofluorescence imaging and fluorescence imaging of two vital dyes: 2-NBDG and proflavine in head and neck cancer detection from surgical specimen. One prominent advantage of HSI is that it does not require the use of an exogenous contrast agent to provide optical contrast, simplifying clinical use. Previously, we have demonstrated the utility of HSI for head and neck cancer detection in subcutaneous cancer animal model [149] [219] [220] as well as a chemically-induced oral cancer model [218]. Uptake of 2-NBDG fluorescent deoxyglucose derivative is associated with increased metabolic activity. A recent study has shown that wide-field optical imaging with 2-NBDG could accurately distinguish the pathologically normal and abnormal biopsies of head and neck cancer patients [221].

Proflavine has also been applied for distinguishing between benign and neoplastic mucosa in the head and neck [222].

Hyperspectral images of resected head and neck tissue were analyzed with two machine learning frameworks: intra-patient and inter-patient classification. Various classifiers have been tested and compared for classifying head and neck tissue. The proposed detection method was validated by a pathological gold standard. Both intra- and inter- patient classification methods demonstrated great potential for evaluating the surgical margins of the *ex vivo* specimen, and for intraoperative surgical guidance. In the intra-patient analysis method, a series of images of the surgical bed before, during, and after tumor resection of the same patient could be acquired for quantitative analysis. Clinically visible tumor and normal tissue spectra may be used to build predictive models and residual tumors in images of surgical bed after resection may be detected with the model. As for the inter-patient analysis method, it would be desirable to establish a spectral database of cancer and normal tissue from a large number of patients and a wide variety of anatomic sites. A predictive model could be built and optimized based on these dataset in advance. The model could be directly used to predict the presence of cancerous tissue during the surgical resection. Furthermore, the model could be utilized to predict the margin status of the resected fresh surgical specimen in the pathology room, which could save the time and cost associated with frozen-section diagnosis while improving accuracy, thus enabling complete resection. Due to large heterogeneity of cancer patients, inter-patient classification was not as good as intra-patient classification in the detection of cancerous tissue in our study. Future work would continue to improve the diagnostic performance for inter-patient classification.

This is only a proof of concept study with a limited number of patients in each anatomic site. We are working on collecting more specimens for different anatomic sites to build a large spectral database for quantitative analysis. In the future, a new HSI instrument needs to be developed and validated for clinical applications. This instrument

should produce high-quality images with minimized glare and be conveniently used in the pathology room and operating rooms. The long-term goal of this work is to provide an adjunct tool for real-time margin assessment during surgery, potentially improving cancer resection while preserving healthy organs.

6.5 Conclusion

In summary, HSI combined with machine learning techniques, enabled the discrimination between normal and cancerous tissue from fresh surgical specimens of a variety of head and neck cancer sites. Although the current study was designed only for imaging the *ex vivo* surgical specimen, the HSI has the potential to be applied for intraoperative tumor margin assessment in the future.

CHAPTER 7

CONCLUSION

7.1 Summary and Research Contributions

HSI has great potential to enable early detection as well as the assessment of surgical margins because of its inherent ability to identify biochemical and morphologic alterations associated with the onset and progression of neoplastic disease. The prominent advantage of HSI is that it combines wide field imaging with spectroscopy. Additionally, HSI is a noninvasive, non-ionizing imaging technology which does not require contrast agents. Hyperspectral images have more spectral channels and higher spectral resolution than RGB images, which can carry more relevant information for characterization of tissue physiology and pathophysiology. Spectroscopy only measures tissue point by point, which might miss regions with the most malignant potential and does not capture spatial patterns. HSI captures the spectral images of a large area of tissue which overcomes the under-sampling problem associated with spectroscopy and conventional biopsy procedures, and can exploit powerful spectral-spatial analysis methods.

Machine learning-based quantitative analysis is critical to exploit the full capability of HSI. In order to make precise diagnosis, it is important to pull out relevant diagnostic information from the enormous volume of spectral-spatial information contained in a hypercube. However, the quantitative analysis of hyperspectral images is challenging due to the large data size, including considerable amounts of spectral redundancy in the highly correlated bands, high dimensionality of spectral bands, and high spatial resolution [189] [223]. Machine learning methods offer powerful quantification tools to mine the rich diagnostic information available from HSI for cancer diagnosis.

This dissertation established the feasibility of HSI combined with a variety of machine learning techniques for the detection and delineation of head and neck cancer in a series of animal experiments and fresh surgical specimens of human patients, which facilitated the future clinical development of HSI technology from bench to bed side.

7.1.1 Review of MHSI

Chapter 2 systematically reviewed the literature on MHSI hardware design, data analysis, and medical applications. This review started at the basics with the mechanisms of HSI and its current development status. Image analysis methods for MHSI were summarized with an emphasis on preprocessing, feature extraction and selection, and classification methods. The section on applications summarized literature on disease diagnosis and surgery guidance. This review provided a comprehensive overview of the literature on MHSI technology and its application.

7.1.2 HSI for noninvasive cancer detection

Chapter 3 described and validated a spectral-spatial classification framework based on tensor modeling for HSI in the application of noninvasive head and neck cancer detection. This tensor-based method proved to be very effective in characterizing both spatial and spectral properties of the hypercube and reducing spectral dimensionality while preserving relevant diagnostic information. The proposed classification method was able to distinguish between tumor and normal tissue in an animal head and neck cancer model. The results from this study demonstrated that the combination of HSI with spectral-spatial classification methods may enable quantitative detection of cancers in a noninvasive manner.

In Chapter 4, another animal experiment was carefully designed to capture hyperspectral images of both *in vivo* and *ex vivo* mouse tongues at multiple time points during tongue carcinogenesis. A chemically-induced tongue carcinogenesis model was

chosen to better simulate the biology and pathology of human carcinogenesis in oral cavity than a subcutaneous cancer model as described in Chapter 3. It became more challenging to detect cancer with HSI in this animal model than in a simple subcutaneous model, due to the small size of the mouse tongues and the difficulty in registration with pathology gold standard. In this Chapter, the pathology gold standard maps were first reconstructed to match the dorsal surface of the tongue for validation. Next, a variety of machine learning classifiers were implemented and validated to compare the diagnostic performance of HSI, autofluorescence imaging (AFI), and fluorescence imaging (FI). Prediction color maps were generated to display the location and distribution of neoplasia. The results demonstrated that HSI outperformed autofluorescence imaging for the detection and delineation of neoplasia in *in vivo* mouse tongues, and the performance of HSI was comparable to AFI, proflavine, and 2-NBDG FI for the detection of *ex vivo* tongue neoplasia. This study demonstrated that HSI holds potential for objective detection and delineation of oral neoplasia. This study demonstrated that *in vivo* HSI can be performed to detect tongue cancer in a mouse model without the use of exogenous contrast agent. New procedures and methods were designed to acquire high quality hyperspectral images of mouse tongue *in vivo* and to align the *in vivo* tongue images acquired with HSI, AF, and FI with *ex vivo* pathology gold standard. The quantification and validation methods could be applied to human tongue cancer detection in the future.

So far we have developed a variety of quantification tools to exploit the whole spectral information provided by HSI. Machine learning techniques provide a fast and accurate approach to predict the disease status of the examined biological tissue for HSI. However, the underlying pathophysiology for the spectral difference between normal and neoplastic tissue is not well understood. So the second part of Chapter 4 developed a novel method to move one step further towards bridging this knowledge gap by linking histological features from digitized histological images quantifying the architectural features of neoplasia on a microscopic scale, with the spectral signature of the

corresponding tissue measured by hyperspectral imaging on a macroscopic level. This work represents the first of its kind for the interpretation of the pathophysiology underlying the hyperspectral dataset.

7.1.3 HSI for surgical margin assessment

HSI is a promising modality for rapid cancer detection during image-guided surgery. But the data from hyperspectral imaging often needs to be processed appropriately in order to extract the maximum useful information that differentiates cancer from normal tissue. In Chapter 5, we proposed a framework for hyperspectral image processing and quantification, which included a set of steps including glare removal, image registration, curvature correction, feature extraction, and ultimately image classification. This was the first time the methods for glare detection and image registration were developed specifically for the pre-processing of hyperspectral images, and these methods can be applicable to other optical images as well. Spectral features including Fourier coefficients, normalized reflectance, mean, and spectral derivatives were analyzed for cancer detection. The experimental results demonstrated the feasibility of the hyperspectral image processing and quantification framework for cancer detection during the animal tumor surgery in a challenging setting where sensitivity can be low due to a modest number of features present but potential for fast image classification can be high. This HSI approach may have the potential application for tumor margin assessment during image-guided surgery, where speed of assessment may be a dominant factor.

Surgical resection is one of the main treatment options for head and neck cancers. The extent of cancer resection is commonly assessed during surgery by pathologic evaluation of the frozen sections at the resected specimen margins to verify whether cancer is present. However, this procedure is time-consuming, labor-intensive, subjective, and suffers from sampling errors. In Chapter 6, a proof-of-concept study was presented to

determine the feasibility of using HSI for identifying and differentiating tumor from normal tissue in fresh surgical specimens from a variety of anatomic sites of head and neck cancer patients. Two analysis frameworks: intra- and inter-patient classification, were proposed to differentiate tumor from normal tissue, both of which have clinical meaningful applications. Cancer probability maps were generated with tumor border outlined, which could provide real-time guidance for tumor resection. The cancerous regions delineated by the automated classification methods were validated by pathological diagnosis. The experimental results suggested that HSI combined with machine learning techniques, could be developed for the assessment of the tumor margins in surgical specimen. This study demonstrated the feasibility of HSI from 450 to 900 nm for the assessment of surgical specimens of head and neck cancer patients.

7.2 Future Directions

As an emerging imaging modality, HSI holds great promise to provide q diagnostic tool for early cancer detection and surgical guidance. However, medical hyperspectral imaging technology is still in its early stage in terms of both the hardware and software development towards clinical applications. In the long run, the development of HSI systems incorporating fast and accurate quantification tools, which are suitable for collecting *in vivo* patient data and validated in large and diverse populations, are crucial to facilitate the clinical translation of HSI technology. Specifically, the following research directions which would facilitate the clinical translation of HSI for the detection and delineation of head and neck cancer from bench to bedside.

Firstly, digging deep into the biophysics of hyperspectral data and identify the key optical parameters associated with the spectral difference between tumor and normal tissue could lay a strong foundation for HSI in cancer research and future translation into the clinics. Development of the inverse modeling methods suitable for hyperspectral data analysis could enable the extraction of the absorption and scattering parameters in two

dimensional images. This type of analysis could directly link these optical parameters with the underlying pathological state, which may better elucidate the rationale of cancer detection with HSI. Moreover, identification of key diagnostic parameters could provide more reliable detection of cancer, and generation of an image with the better tumor-normal contrast could assist the clinicians on decision making.

Secondly, a portable HSI system needs to be developed for early cancer detection in human patients. The HSI system used in this dissertation is a small animal imaging system, which is limited to imaging small animal models *in vivo* or tissue specimens *ex vivo*. With a portable imaging system, it is possible to directly scan the entire human oral cavity noninvasively and potentially capture the premalignant lesions such as leukoplakia, and erythroplakia, which are not available in mouse models. Clinical studies can be designed to further assess the diagnostic sensitivity and specificity of HSI for the detection of pre-malignant lesions in a large and diverse population. Furthermore, by incorporation of this system with endoscope, more head and neck sites beyond oral cavity, such as esophagus, throat, and nasal cavity, could be examined with improved accuracy.

Thirdly, a new HSI system suitable for rapid cancer margin assessment in both the pathology room and the operating room could be developed. Currently, the size of the collected surgical specimen is relatively small, due to the limited availability of human tissue for research. If HSI could be designed for use in pathology rooms, the entire surgical specimen could be imaged and the surgical margins could be directly examined with HSI technology. This could provide a fast diagnostic aid for surgical pathologist and potentially save the time and costs. To achieve this, more intensive validation of HSI on heterogeneous tissue samples of a larger population need to be conducted.

The application of HSI for intraoperative surgical guidance still has a long way to go, with many challenging problems to be addressed. For example, ambient light could greatly affect the reflectance signal by HSI. Currently, all the imaging experiments were

conducted inside a dark imaging box without the effect of the ambient light, which is not the case in the surgical room with multiple light sources, such as the surgical light, room light, and light from the monitoring devices. Understanding the characteristics of these light sources would be helpful to develop solutions to alleviate the effect of these light sources in the surgical room. One potential solution is to modulate the light source of HSI with certain frequencies and extract the reflectance signal during post-processing. Furthermore, blooding may be a problem to acquire clean and consistent reflectance signal with HSI during the surgery. Clean up the surgical field quickly right before the imaging could be helpful. Development of quantification methods to extract the reflectance signal buried by the blood would also be valuable.

In terms of the system design, two aspects should be considered: (1) The NIR wavelength region above 900 nm could be included and explored to enhance the noninvasive detection of deeply resided tumors, because it can penetrate deeper into the tissue and provide complementary information for cancer diagnostics. (2) Address the glares commonly seen during intraoperative imaging, which is caused by specular reflection of the liquid on the tissue surface, deteriorating the quality of imaging and affecting the quantitative analysis of the images. Therefore, it would be beneficial to consider the elimination or reduction of glare during the design of the new instrument. One potential solution is to add cross-polarization by placing a linear polarizer in front of the light source and an orthogonal polarizer in front of the detector. The orthogonal orientation of the second polarizer has the effect of selectively passing through photons which have undergone multiple scattering events in the tissue and rejecting the photons from specular reflection from tissue surface.

As for software development, quantification tools need to be implemented and incorporated into the HSI systems. These systems should be validated in large patient population to prove its accuracy and robustness and to provide real-time cancer diagnosis in the clinics. More specifically, a large spectral library could be created by collecting

hyperspectral images of different anatomic sites, as well as different tissue components such as epithelium, connective tissue, skeletal muscle, and fat from large populations. Analysis of intra- and inter-patient spectral variations would be important in order to provide robust and consistent diagnosis by HSI. Machine learning algorithms incorporating both spectral-spatial image features and pixel-wise optical parameters may provide complementary information and boost the diagnostic performance of HSI. Deep learning methods which can automatically learn the underlying patterns of data could also be explored to mine the large volume of hyperspectral data.

Moving forward, the medical applications of HSI could be extended beyond the detection and surgical guidance of head and neck cancer, by incorporating HSI systems with conventional optical imaging techniques such as endoscopy, colpscope, laparoscope, fundus camera, and microscope.

In summary, the biophysics study of the hyperspectral data, development of HSI systems suitable for clinical applications, and the fast and accurate quantification tools validated intensively in large patient populations, would be the key steps towards the clinical translation of HSI technology, and to make the broader optical community and medical community to adopt and embrace this technology.

APPENDIX A

SELECTED PUBLICATIONS

The work discussed in this thesis is a compilation of several years of research that resulted in the following peer-reviewed journal publications, and conference proceedings.

Journal Publications

1. **G Lu**, B Fei. "Medical Hyperspectral Imaging: A Review," *Journal of Biomedical Optics*. 19(1):10901. 2014.

2. **G Lu**, L Halig, D Wang, X Qin, ZG Chen, B Fei. "Spatial-Spectral Classification for Noninvasive Cancer Detection Using Hyperspectral Imaging," *Journal of Biomedical Optics* 19(10), 106004. 2014.

3. **G Lu**, D Wang, X Qin, L Halig, S Muller, H Zhang, A Chen, BW Pogue, ZG Chen, B Fei. "Framework for Hyperspectral Image Processing and Quantification for Cancer Detection during Animal Tumor Surgery," *Journal of Biomedical Optics* 20(12), 126012. 2015.

4. R Pike, **G Lu**[*], D Wang, ZG Chen, B Fei. "A Minimum Spanning Forest Based Method for Detection of Cancerous Tissue with Hyperspectral Imaging," *IEEE Transactions on Biomedical Engineering*. vol. PP, no. 99, pp.1-1 2015. [*]**Co-first author**.

Manuscripts Submitted and In Preparation

5. **G Lu**, D Wang, X Qin, S Muller, Z Chen, A Chen, B Fei. " Detection and delineation of squamous neoplasia with hyperspectral imaging in a mouse model of tongue carcinogenesis,"*Journal of Biophotonics*, Submitted.

6. **G Lu**, D Wang, X Qin, S Muller, Z Chen, A Chen, B Fei. " Histopathology Feature Mining and Association with Hyperspectral Imaging for the Detection of Squamous Neoplasia," *Scientific Report*, Submitted.

7. **G Lu**, H Zhang, X Wang, J Little, K Magliocca, A Chen, B Fei. " Quantitative Detection of Head and Neck Cancer in Fresh Surgical Specimen of Human Patients Using Hyperspectral Imaging," *Clinical Cancer Research*, Submitted.

Conference Proceedings

1. **G Lu**, X Qin, D Wang, S Muller, H Zhang, A Chen, ZG Chen, B Fei. "Quantitative Tongue Cancer Diagnosis with Image Classification on Histological Images," *Proc. SPIE 9791, Medical Imaging 2016: Digital Pathology*, 97910L.

2. **G Lu**, X Qin, D Wang, S Muller, H Zhang, A Chen, ZG Chen, B Fei. "Hyperspectral Imaging of Neoplastic Progression in a Mouse Model of Oral Carcinogenesis," *Proc. SPIE 9788, Medical Imaging 2016: Biomedical Applications in Molecular, Structural, and Functional Imaging*, 978812.

3. H Chung, **G Lu**, Z Tian, D Wang, ZG Chen, B Fei. "Superpixel-based spectral classification for the detection of head and neck cancer with hyperspectral imaging," Proc. *SPIE 9788, Medical Imaging 2016: Biomedical Applications in Molecular, Structural, and Functional Imaging*, 978813.

4. **G Lu**, X Qin, D Wang, ZG Chen, B Fei. "Quantitative Wavelength Analysis and Image Classification for Intraoperative Cancer Diagnosis with Hyperspectral Imaging," *Proc. SPIE. 9415, Medical Imaging 2015: Image-Guided Procedures, Robotic Interventions, and Modeling*, 94151B.

5. **G Lu**, X Qin, D Wang, ZG Chen, B Fei. "Estimation of Tissue Optical Parameters with Hyperspectral Imaging and Spectral Unmixing," *Proc. SPIE 9417, Medical Imaging 2015: Biomedical Applications in Molecular, Structural, and Functional Imaging*, 94170Q.

6. **G Lu**, L Halig, D Wang, ZG Chen, B Fei. "Spectral-Spatial Classification Using Tensor Modeling for Head and Neck Cancer Detection Using Hyperspectral Imaging," Proc. SPIE 9034, Medical Imaging 2014: Image Processing, 903413.

7. B Fei, **G Lu**, R Pike, D Wang, ZG Chen. "WE-D-9A-05: Medical Hyperspectral Imaging for the Detection of Head and Neck Cancer in Animal Models." Medical Physics 41.6 (2014): 500-501.

8. **G Lu**, L Halig, D Wang, ZG Chen, B Fei. "Hyperspectral Imaging for Surgical Margin Delineation of Head and Neck Cancer: Registration of Hyperspectral and Histological Images," *Proc. SPIE. 9036, Medical Imaging 2014: Image-Guided Procedures, Robotic Interventions, and Modeling*, 90360S.

9. R Pike, SK Patton, **G Lu**, L Halig, D Wang, ZG Chen, B Fei. "A minimum spanning forest based hyperspectral image classification method for cancerous tissue detection," *Proc. SPIE. 9034, Medical Imaging 2014: Image Processing*, 90341W.

REFERENCES

[1] E. Wissinger, I. Griebsch, J. Lungershausen, T. Foster, and C. Pashos, "The Economic Burden of Head and Neck Cancer: A Systematic Literature Review," *PharmacoEconomics,* vol. 32, pp. 865-882, 2014.

[2] S. Marur and A. A. Forastiere, "Head and Neck Squamous Cell Carcinoma: Update on Epidemiology, Diagnosis, and Treatment," *Mayo Clinic Proceedings*, vol. 91, pp. 386-396, 2016.

[3] L. A. Torre, F. Bray, R. L. Siegel, J. Ferlay, J. Lortet-Tieulent, and A. Jemal, "Global cancer statistics, 2012," *CA: A Cancer Journal for Clinicians,* vol. 65, pp. 87-108, 2015.

[4] R. I. Haddad and D. M. Shin "Recent Advances in Head and Neck Cancer," *New England Journal of Medicine,* vol. 359, pp. 1143-1154, 2008.

[5] E. Surveillance, and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database. (March 19). *SEER 18 2005-2011, All Races, Both Sexes by SEER Summary Stage 2000. National Cancer Institute, DCCPS, Surveillance Research Program, Surveillance Systems Branch*.

[6] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2016," *CA: A Cancer Journal for Clinicians*, vol. 66, pp. 7-30, 2016.

[7] T. J. Smith and B. E. Hillner, "Bending the Cost Curve in Cancer Care," *New England Journal of Medicine*, vol. 364, pp. 2060-2065, 2011.

[8] R. Etzioni, N. Urban, S. Ramsey, M. McIntosh, S. Schwartz, B. Reid, *et al.*, "The case for early detection," *Nat Rev Cancer,* vol. 3, pp. 243-52, 2003.

[9] N. Bedard, M. Pierce, A. El-Nagger, S. Anandasabapathy, A. Gillenwater, and R. Richards-Kortum, "Emerging roles for multimodal optical imaging in early cancer detection: a global challenge," *Technol Cancer Res Treat*, vol. 9, pp. 211-7, 2010.

[10] M. W. M. van den Brekel, W. L. Lodder, H. V. Stel, E. Bloemena, C. R. Leemans, and I. van der Waal, "Observer variation in the histopathologic

assessment of extranodal tumor spread in lymph node metastases in the neck," *Head & Neck,* vol. 34, pp. 840-845, 2012.

[11] D. Roblyer, R. Richards-Kortum, K. Sokolov, A. K. El-Naggar, M. D. Williams, C. Kurachi, *et al.*, "Multispectral optical imaging device for in vivo detection of oral neoplasia," *Journal of biomedical optics,* vol. 13, p. 024019, 2008.

[12] S. M. Ismail, A. B. Colclough, J. S. Dinnen, D. Eakins, D. M. Evans, E. Gradwell, *et al.*, "Observer variation in histopathological diagnosis and grading of cervical intraepithelial neoplasia," *Bmj,* vol. 298, pp. 707-10, 1989.

[13] A. Argiris, M. V. Karamouzis, D. Raben, and R. L. Ferris, "Head and neck cancer," *The Lancet,* vol. 371, pp. 1695-1709, 2008.

[14] Q. T. Nguyen and R. Y. Tsien, "Fluorescence-guided surgery with live molecular navigation - a new cutting edge," *Nat Rev Cancer,* vol. 13, pp. 653-62, 2013.

[15] Q. T. Nguyen, E. S. Olson, T. A. Aguilera, T. Jiang, M. Scadeng, L. G. Ellies, *et al.*, "Surgery with molecular fluorescence imaging using activatable cell-penetrating peptides decreases residual cancer and improves survival," *Proceedings of the National Academy of Sciences,* vol. 107, pp. 4317-4322, 2010.

[16] S. Keereweer, J. F. Kerrebijn, P. A. A. Driel, B. Xie, E. Kaijzel, T. A. Snoeks, *et al.*, "Optical Image-guided Surgery—Where Do We Stand?," *Molecular Imaging and Biology,* vol. 13, pp. 199-207, 2011.

[17] R. F. Gandour-Edwards, P. J. Donald, and D. A. Wiese, "Accuracy of intraoperative frozen section diagnosis in head and neck surgery: experience at a university medical center," *Head Neck,* vol. 15, pp. 33-8, 1993.

[18] A. F. H. Goetz, "Three decades of hyperspectral remote sensing of the Earth: A personal view," *Remote Sensing of Environment,* vol. 113, Supplement 1, pp. S5-S16, 2009.

[19] C. Fischer and I. Kakoulli, "Multispectral and hyperspectral imaging technologies in conservation: current research and potential applications," *Studies in Conservation,* vol. 51, pp. 3-16, 2006.

[20] H. Liang, "Advances in multispectral and hyperspectral imaging for archaeology and art conservation," *Applied Physics A*, vol. 106, pp. 309-323, 2012.

[21] M. Govender, K. Chetty, and H. Bulcock, "A review of hyperspectral remote sensing and its application in vegetation and water resource studies," *Water SA*, vol. 33, pp. 145-151, 2007.

[22] E. Adam, O. Mutanga, and D. Rugege, "Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review," *Wetlands Ecology and Management*, vol. 18, pp. 281-296, 2010.

[23] A. A. Gowen, C. P. O'Donnell, P. J. Cullen, G. Downey, and J. M. Frias, "Hyperspectral imaging – an emerging process analytical tool for food quality and safety control," *Trends in Food Science & Technology*, vol. 18, pp. 590-598, 2007.

[24] Y. Z. Feng and D. W. Sun, "Application of hyperspectral imaging in food safety inspection and control: a review," *Crit Rev Food Sci Nutr*, vol. 52, pp. 1039-58, 2012.

[25] G. J. Edelman, E. Gaston, T. G. van Leeuwen, P. J. Cullen, and M. C. G. Aalders, "Hyperspectral imaging for non-contact analysis of forensic traces," *Forensic Science International*, vol. 223, pp. 28-39, 2012.

[26] D. B. Malkoff and W. R. Oliver, "Hyperspectral imaging applied to forensic medicine," in *Proc. SPIE 3920, Spectral Imaging: Instrumentation, Applications, and Analysis*, 2000, pp. 108-116.

[27] J. Kuula, I. Pölönen, H.-H. Puupponen, T. Selander, T. Reinikainen, T. Kalenius, *et al.*, "Using VIS/NIR and IR spectral cameras for detecting and separating crime scene details," in *Proc. SPIE 8359, Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense XI*, 2012, p. 83590P.

[28] R. L. Schuler, P. E. Kish, and C. A. Plese, "Preliminary Observations on the Ability of Hyperspectral Imaging to Provide Detection and Visualization of Bloodstain Patterns on Black Fabrics," *Journal of Forensic Sciences*, vol. 57, pp. 1562-1569, 2012.

[29] O. Carrasco, R. B. Gomez, A. Chainani, and W. E. Roper, "Hyperspectral imaging applied to medical diagnoses and food safety," in *Proc. SPIE 5097, Geo-Spatial and Temporal Image and Data Exploitation III*, 2003, pp. 215-221.

[30] M. A. Afromowitz, J. B. Callis, D. M. Heimbach, L. A. DeSoto, and M. K. Norton, "Multispectral imaging of burn wounds: a new clinical instrument for evaluating burn depth," *Biomedical Engineering, IEEE Transactions on,* vol. 35, pp. 842-850, 1988.

[31] G. Zonios, L. T. Perelman, V. Backman, R. Manoharan, M. Fitzmaurice, J. Van Dam, *et al.*, "Diffuse reflectance spectroscopy of human adenomatous colon polyps in vivo," *Applied Optics,* vol. 38, pp. 6628-6637, 1999.

[32] L. V. Wang and H.-I. Wu, "Introduction," in *Biomedical Optics*, ed: John Wiley & Sons, Inc., 2009, pp. 1-15.

[33] B. Costas, P. Christos, and E. George, "Multi/Hyper-Spectral Imaging," in *Handbook of Biomedical Optics*, ed: CRC Press, 2011, pp. 131-164.

[34] V. V. Tuchin and V. Tuchin, *Tissue optics: light scattering methods and instruments for medical diagnosis* vol. 13: SPIE press Bellingham, 2007.

[35] D. G. Ferris, R. A. Lawhead, E. D. Dickman, N. Holtzapple, J. A. Miller, S. Grogan, *et al.*, "Multimodal Hyperspectral Imaging for the Noninvasive Diagnosis of Cervical Neoplasia," *Journal of Lower Genital Tract Disease,* vol. 5, pp. 65-72, 2001.

[36] M. C. Pierce, R. A. Schwarz, V. S. Bhattar, S. Mondrik, M. D. Williams, J. J. Lee, *et al.*, "Accuracy of In Vivo Multimodal Optical Imaging for Detection of Oral Neoplasia," *Cancer Prevention Research,* vol. 5, pp. 801-809, 2012.

[37] G. Lu and B. Fei, "Medical hyperspectral imaging: a review," *J Biomed Opt,* vol. 19, p. 10901, 2014.

[38] A. Rogalski, "Progress in focal plane array technologies," *Progress in Quantum Electronics,* vol. 36, pp. 342-473, 2012.

[39] K. J. Zuzak, M. D. Schaeberle, I. W. Levin, N. E. Lewis, J. Freeman, J. D. McNeil, *et al.*, "Visible and infrared hyperspectral visualization of normal and

ischemic tissue," in *Engineering in Medicine and Biology, 1999. BMES/EMBS Conference, 1999. Proceedings of the First Joint*, 1999, p. 1118.

[40] K. J. Zuzak, M. D. Schaeberle, E. N. Lewis, and I. W. Levin, "Visible spectroscopic imaging studies of normal and ischemic dermal tissue," in *Proc. SPIE 3918, Biomedical Spectroscopy: Vibrational Spectroscopy and Other Novel Techniques*, 2000, pp. 17-26.

[41] K. J. Zuzak, M. D. Schaeberle, E. N. Lewis, and I. W. Levin, "Visible Reflectance Hyperspectral Imaging: Characterization of a Noninvasive, in Vivo System for Determining Tissue Perfusion," *Analytical Chemistry,* vol. 74, pp. 2021-2028, 2002.

[42] L. L. Randeberg, I. Baarstad, T. Løke, P. Kaspersen, and L. O. Svaasand, "Hyperspectral imaging of bruised skin," in *Proc. SPIE 6078, Photonic Therapeutics and Diagnostics II*, 2006, p. 60780O.

[43] E. L. P. Larsen, L. L. Randeberg, E. Olstad, O. A. Haugen, A. Aksnes, and L. O. Svaasand, "Hyperspectral imaging of atherosclerotic plaques in vitro," *Journal of Biomedical Optics,* vol. 16, pp. 026011-026011, 2011.

[44] M. Patterson, B. Wilson, and D. Wyman, "The propagation of optical radiation in tissue I. Models of radiation transport and their application," *Lasers in Medical Science,* vol. 6, pp. 155-168, 1991.

[45] M. Joel and V.-D. Tuan, "Optical Properties of Tissue," in *Biomedical Photonics Handbook*, ed: CRC Press, 2003, pp. 1-76.

[46] Y. Zhang, Y. Chen, Y. Yu, X. Xue, V. V. Tuchin, and D. Zhu, "Visible and near-infrared spectroscopy for distinguishing malignant tumor tissue from benign tumor and normal breast tissues in vitro," *J Biomed Opt,* vol. 18, p. 077003, 2013.

[47] A. Welch, M. C. Gemert, and W. Star, "Definitions and Overview of Tissue Optics," in *Optical-Thermal Response of Laser-Irradiated Tissue*, A. J. Welch and M. J. C. Gemert, Eds., ed: Springer Netherlands, 2011, pp. 27-64.

[48] A. Roggan, K. Dörschel, O. Minet, D. Wolff, and G. Müller, "The optical properties of biological tissue in the near infrared wavelength range: review and measurements," in *LITT: laser-induced interstitial thermotherapy. Workshop*, 1995, pp. 10-44.

[49] J. A. Freeberg, J. L. Benedet, C. MacAulay, L. A. West, and M. Follen, "The performance of fluorescence and reflectance spectroscopy for the in vivo diagnosis of cervical neoplasia; point probe versus multispectral approaches," *Gynecologic Oncology,* vol. 107, pp. S248-S255, 2007.

[50] V.-D. Tuan and M. C. Brian, "Fluorescence Spectroscopy for Biomedical Diagnostics," in *Biomedical Photonics Handbook*, ed: CRC Press, 2003, pp. 1-51.

[51] V. T. Valery, "Light-Tissue Interactions," in *Biomedical Photonics Handbook*, ed: CRC Press, 2003, pp. 1-27.

[52] M. Sasha, M. Jelena, and F. Michael, "Reflectance Spectroscopy," in *Handbook of Biomedical Optics*, ed: CRC Press, 2011, pp. 103-130.

[53] K. Sokolov, M. Follen, and R. Richards-Kortum, "Optical spectroscopy for detection of neoplasia," *Current Opinion in Chemical Biology,* vol. 6, pp. 651-658, 2002.

[54] D. Landgrebe, "Hyperspectral image data analysis," *Signal Processing Magazine, IEEE,* vol. 19, pp. 17-28, 2002.

[55] S. T. Monteiro, Y. Kosugi, K. Uto, and E. Watanabe, "Towards Applying hyperspectral Imagery as an Intraoperative Visual Aid Tool," in *Proc. 4th International Conference on Visualization, Imaging and Image Processing*, 2004, pp. 483-488.

[56] Z. Liu, H. J. Wang, and Q. L. Li, "Tongue Tumor Detection in Medical Hyperspectral Images," *Sensors,* vol. 12, pp. 162-174, 2012.

[57] H. Akbari, Y. Kosugi, K. Kojima, and N. Tanaka, "Detection and Analysis of the Intestinal Ischemia Using Visible and Invisible Hyperspectral Imaging," *Biomedical Engineering, IEEE Transactions on,* vol. 57, pp. 2011-2017, 2010.

[58] S. G. Kong, Z. Du, M. Martin, and T. Vo-Dinh, "Hyperspectral fluorescence image analysis for use in medical diagnostics," in *Advanced Biomedical and Clinical Diagnostic Systems III*. vol. 5692, ed, 2005, pp. 21-28.

[59] A. Sharwani, W. Jerjes, V. Salih, B. Swinson, I. J. Bigio, M. El-Maaytah, *et al.*, "Assessment of oral premalignancy using elastic scattering spectroscopy," *Oral Oncol,* vol. 42, pp. 343-9, Apr 2006.

[60] S. V. Panasyuk, S. Yang, D. V. Faller, D. Ngo, R. A. Lew, J. E. Freeman, *et al.*, "Medical hyperspectral imaging to facilitate residual tumor identification during surgery," *Cancer Biology & Therapy,* vol. 6, pp. 439-446, 2007.

[61] A. Plaza, J. A. Benediktsson, J. W. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, *et al.*, "Recent advances in techniques for hyperspectral image processing," *Remote Sensing of Environment,* vol. 113, Supplement 1, pp. S110-S122, 2009.

[62] M. Fauvel, Y. Tarabalka, J. n. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Advances in Spectral–Spatial Classification of Hyperspectral Images," *Proceedings of the IEEE,* 2012.

[63] X. Zhang and C. H. Chen, "New independent component analysis method using higher order statistics with application to remote sensing images," *Optical Engineering,* vol. 41, pp. 1717-1728, 2002.

[64] R. A. Fisher, "THE USE OF MULTIPLE MEASUREMENTS IN TAXONOMIC PROBLEMS," *Annals of Eugenics,* vol. 7, pp. 179-188, 1936.

[65] Y. Guo, T. Hastie, and R. Tibshirani, "Regularized linear discriminant analysis and its application in microarrays," *Biostatistics,* 2006.

[66] D. Landgrebe, "Hyperspectral image data analysis," *IEEE Signal Processing Magazine,* vol. 19, pp. 17-28, 2002.

[67] T. Li, S. Zhu, and M. Ogihara, "Using discriminant analysis for multi-class classification: an experimental investigation," *Knowledge and Information Systems,* vol. 10, pp. 453-472, 2006.

[68] A. Shashua, "On the Relationship Between the Support Vector Machine for Classification and Sparsified Fisher's Linear Discriminant," *Neural Processing Letters,* vol. 9, pp. 129-139, 1999.

[69] J. J. Squiers, W. Li, D. R. King, W. Mo, X. Zhang, Y. Lu, *et al.*, "Multispectral imaging burn wound tissue classification system: a comparison of test accuracies

between several common machine learning algorithms," in *Proc. SPIE 9785, Medical Imaging 2016*, 2016, pp. 97853L-97853L-10.

[70] H. Tin Kam, "The random subspace method for constructing decision forests," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 20, pp. 832-844, 1998.

[71] C. Seiffert, T. M. Khoshgoftaar, J. V. Hulse, and A. Napolitano, "RUSBoost: A Hybrid Approach to Alleviating Class Imbalance," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans,* vol. 40, pp. 185-197, 2010.

[72] L. Breiman, "Random Forests," *Machine Learning,* vol. 45, pp. 5-32, 2001.

[73] M. Belgiu and L. Drăguţ, "Random forest in remote sensing: A review of applications and future directions," *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 114, pp. 24-31, 2016.

[74] G. Lu, X. Qin, D. Wang, S. Muller, H. Zhang, A. Chen, *et al.*, "Hyperspectral imaging of neoplastic progression in a mouse model of oral carcinogenesis," in *Proc. SPIE 9788*, 2016, pp. 978812-978812-8.

[75] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *Geoscience and Remote Sensing, IEEE Transactions on,* vol. 42, pp. 1778-1790, 2004.

[76] Z. Liu, J. Q. Yan, D. Zhang, and Q. L. Li, "Automated tongue segmentation in hyperspectral images for medicine," *Appl Opt,* vol. 46, pp. 8328-34, 2007.

[77] Z. Liu, D. Zhang, J. Q. Yan, Q. L. Li, and Q. L. Tang, "Classification of hyperspectral medical tongue images for tongue diagnosis," *Computerized Medical Imaging and Graphics,* vol. 31, pp. 672-678, 2007.

[78] H. Akbari, L. V. Halig, D. M. Schuster, A. Osunkoya, V. Master, P. T. Nieh, *et al.*, "Hyperspectral imaging and quantitative analysis for prostate cancer detection," *Journal of biomedical optics,* vol. 17, p. 076005, 2012.

[79] H. Akbari, L. V. Halig, H. Zhang, D. Wang, Z. G. Chen, and B. Fei, "Detection of Cancer Metastasis Using a Novel Macroscopic Hyperspectral Method," in *Proc.*

SPIE 8317, Medical Imaging 2012: Biomedical Applications in Molecular, Structural, and Functional Imaging, 2012, p. 831711.

[80] H. Akbari, Y. Kosugi, K. Kojima, and N. Tanaka, "Blood vessel detection and artery-vein differentiation using hyperspectral imaging," in *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, 2009, pp. 1461-1464.

[81] H. Akbari, K. Uto, Y. Kosugi, K. Kojima, and N. Tanaka, "Cancer detection using infrared hyperspectral imaging," *Cancer Science*, vol. 102, pp. 852-857, 2011.

[82] S. Kong and L.-J. Park, "Hyperspectral Image Analysis for Skin Tumor Detection," in *Augmented Vision Perception in Infrared*, R. Hammoud, Ed., ed, 2009, pp. 155-171.

[83] K. Masood and N. Rajpoot, "Texture based classification of hyperspectral colon biopsy samples using CLBP," in *Biomedical Imaging: From Nano to Macro, 2009. ISBI '09. IEEE International Symposium on*, 2009, pp. 1011-1014.

[84] K. Rajpoot and N. Rajpoot, "SVM Optimization for Hyperspectral Colon Tissue Cell Classification," in *Medical Image Computing and Computer-Assisted Intervention-MICCAI* vol. 3217, ed: Springer Berlin Heidelberg, 2004, pp. 829-837.

[85] M. Fauvel, J. A. Benediktsson, J. Chanussot, and J. R. Sveinsson, "Spectral and Spatial Classification of Hyperspectral Data Using SVMs and Morphological Profiles," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 46, pp. 3804-3814, 2008.

[86] G. Hughes, "On the mean accuracy of statistical pattern recognizers," *Information Theory, IEEE Transactions on*, vol. 14, pp. 55-63, 1968.

[87] C. Cortes and V. Vapnik, "Support vector machine," *Machine learning*, vol. 20, pp. 273-297, 1995.

[88] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, pp. 1-27, 2011.

[89] D. Lu and Q. Weng, "A survey of image classification methods and techniques for improving classification performance," *International Journal of Remote Sensing*, vol. 28, pp. 823-870, 2007.

[90] F. Blanco, M. López-Mesas, S. Serranti, G. Bonifazi, J. Havel, and M. Valiente, "Hyperspectral imaging based method for fast characterization of kidney stone types," *Journal of Biomedical Optics*, vol. 17, pp. 076027-1, 2012.

[91] R. Jolivot, P. Vabres, and F. Marzani, "Reconstruction of hyperspectral cutaneous data from an artificial neural network-based multispectral imaging system," *Computerized Medical Imaging and Graphics*, vol. 35, pp. 85-88, 2011.

[92] H. Akbari, Y. Kosugi, K. Kojima, and N. Tanaka, "Wavelet-Based Compression and Segmentation of Hyperspectral Images in Surgery," in *Medical Imaging and Augmented Reality*. vol. 5128, ed, 2008, pp. 142-149.

[93] J. M. Benavides, S. Chang, S. Y. Park, R. Richards-Kortum, N. Mackinnon, C. MacAulay, *et al.*, "Multispectral digital colposcopy for in vivo detection of cervical cancer," *Optics Express*, vol. 11, pp. 1223-1236, 2003.

[94] A. O. N. Joseph, "Hyperspectral optical imaging for detection, diagnosis and staging of cancer," *University of Southern California, PhD Dissertation*, 2012.

[95] L. C. Cancio, A. I. Batchinsky, J. R. Mansfield, S. Panasyuk, K. Hetz, D. Martini, *et al.*, "Hyperspectral Imaging: A New Approach to the Diagnosis of Hemorrhagic Shock," *The Journal of Trauma and Acute Care Surgery*, vol. 60, pp. 1087-1095, 2006.

[96] L. C. Cancio, "Application of novel hyperspectral imaging technologies in combat casualty care," in *Proc. SPIE 7596, Emerging Digital Micromirror Device Based Systems and Applications II*, 2010, p. 759605.

[97] J. A. Chin, E. C. Wang, and M. R. Kibbe, "Evaluation of hyperspectral technology for assessing the presence and severity of peripheral artery disease," *Journal of Vascular Surgery*, vol. 54, pp. 1679-1688, 2011.

[98] P. Usenik, M. Bürmen, A. Fidler, F. Pernuš, and B. Likar, "Evaluation of cross-polarized near infrared hyperspectral imaging for early detection of dental caries," in *Proc. SPIE 8208, Lasers in Dentistry XVIII*, 2012, p. 82080G.

[99] R. Martin, B. Thies, and A. Gerstner, "Hyperspectral hybrid method classification for detecting altered mucosa of the human larynx," *International Journal of Health Geographics,* vol. 11, p. 21, 2012.

[100] L. V. Wang and H.-i. Wu, *Biomedical optics: principles and imaging*: John Wiley & Sons, 2012.

[101] K. Sokolov, M. Follen, and R. Richards-Kortum, "Optical spectroscopy for detection of neoplasia," *Curr Opin Chem Biol,* vol. 6, pp. 651-658, 2002.

[102] A. M. Siddiqi, H. Li, F. Faruque, W. Williams, K. Lai, M. Hughson, *et al.,* "Use of hyperspectral imaging to distinguish normal, precancerous, and cancerous cells," *Cancer Cytopathology,* vol. 114, pp. 13-21, 2008.

[103] I. Pavlova, K. Sokolov, R. Drezek, A. Malpica, M. Follen, and R. Richards-Kortum, "Microanatomical and biochemical origins of normal and precancerous cervical autofluorescence using laser-scanning fluorescence confocal microscopy," *Photochem Photobiol,* vol. 77, pp. 550-5, 2003.

[104] L. Boucheron, Z. Bi, N. Harvey, B. Manjunath, and D. Rimm, "Utility of multispectral imaging for nuclear classification of routine clinical histopathology imagery," *BMC Cell Biol,* vol. 8, p. S8, 2007.

[105] S. Argov, J. Ramesh, I. Sinelnikov, H. Guterman, S. Mordechai, J. Goldstein, *et al.*, "Diagnostic potential of Fourier-transform infrared microspectroscopy and advanced computational methods in colon cancer patients," *Journal of Biomedical Optics,* vol. 7, pp. 248-254, 2002.

[106] M. Maggioni, G. L. Davis, F. J. Warner, F. B. Geshwind, A. C. Coppi, R. A. DeVerse, *et al.*, "Hyperspectral microscopic analysis of normal, benign and carcinoma microarray tissue sections," in *Proc. SPIE 6091, Optical Biopsy VI*, 2006, p. 60910I.

[107] K. M. Rajpoot and N. M. Rajpoot, "Wavelet based segmentation of hyperspectral colon tissue imagery," in *Multi Topic Conference, 2003. INMIC 2003. 7th International*, 2003, pp. 38-43.

[108] K. Rajpoot and N. Rajpoot, "SVM Optimization for Hyperspectral Colon Tissue Cell Classification," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI* vol. 3217, ed, 2004, pp. 829-837.

[109] K. Masood, N. Rajpoot, K. Rajpoot, and H. Qureshi, "Hyperspectral Colon Tissue Classification using Morphological Analysis," in *Emerging Technologies, 2006. ICET '06. International Conference on*, 2006, pp. 735-741.

[110] K. Masood and N. M. Rajpoot, "Classification of Colon Biopsy Samples by Spatial Analysis of a Single Spectral Band from its Hyperspectral Cube," in *Proceedings Medical Image Understanding and Analysis (MIUA)*, 2007, pp. 42–48.

[111] K. Masood, "Hyperspectral imaging with wavelet transform for classification of colon tissue biopsy samples," in *Proc. SPIE 7073, Applications of Digital Image Processing XXXI*, 2008, p. 707319.

[112] K. Masood and N. Rajpoot, "Spatial Analysis for Colon Biopsy Classification from Hyperspectral Imagery," *Annals of BMVA*, vol. 2008, pp. 1-16, 2008.

[113] M. Hohmann, A. Douplik, J. Varadhachari, A. Nasution, J. Mudter, M. Neurath, *et al.*, "Preliminary results for hyperspectral videoendoscopy diagnostics on the phantoms of normal and abnormal tissues: towards gastrointestinal diagnostics," in *Proc. SPIE 8087, Clinical and Biomedical Spectroscopy and Imaging II*, 2011, p. 80872N.

[114] S. Kiyotoki, J. Nishikawa, T. Okamoto, K. Hamabe, M. Saito, A. Goto, *et al.*, "New method for detection of gastric cancer by hyperspectral imaging: a pilot study," *J Biomed Opt*, vol. 18, pp. 026010-026010, 2013.

[115] D. T. Dicker, J. Lerner, P. Van Belle, D. Guerry, M. Herlyn, D. E. Elder, *et al.*, "Differentiation of normal skin and melanoma using high resolution hyperspectral imaging," *Cancer Biology & Therapy*, vol. 5, pp. 1033-1038, 2006.

[116] D. Hattery, M. Hassan, S. Demos, and A. Gandjbakhche, "Hyperspectral imaging of Kaposi's Sarcoma for disease assessment and treatment monitoring," in *Applied Imagery Pattern Recognition Workshop, 2002. Proceedings. 31st*, 2002, pp. 124-130.

[117] T. E. Renkoski, K. D. Hatch, and U. Utzinger, "Wide-field spectral imaging of human ovary autofluorescence and oncologic diagnosis via previously collected probe data," *Journal of Biomedical Optics*, vol. 17, pp. 036003-1, 2012.

[118] C. Angeletti, N. R. Harvey, V. Khomitch, A. H. Fischer, R. M. Levenson, and D. L. Rimm, "Detection of malignancy in cytology specimens using spectral–spatial analysis," *Laboratory investigation,* vol. 85, pp. 1555-1564, 2005.

[119] M. E. Martin, M. B. Wabuyele, M. Panjehpour, M. N. Phan, B. F. Overholt, and T. Vo-Dinh, "Hyperspectral fluorescence imaging system for biomedical diagnostics," in *Proc. SPIE 6080, Advanced Biomedical and Clinical Diagnostic Systems IV*, 2006, p. 60800Q.

[120] M. E. Martin, M. B. Wabuyele, K. Chen, P. Kasili, M. Panjehpour, M. Phan, *et al.*, "Development of an Advanced Hyperspectral Imaging (HSI) system with application for cancer detection," *Annals of Biomedical Engineering,* vol. 34, pp. 1061-1068, 2006.

[121] D. Roblyer, R. Richards-Kortum, K. Sokolov, A. K. El-Naggar, M. D. Williams, C. Kurachi, *et al.*, "Multispectral optical imaging device for in vivo detection of oral neoplasia," *J Biomed Opt,* vol. 13, pp. 024019-024019, 2008.

[122] D. Roblyer, C. Kurachi, A. M. Gillenwater, and R. Richards-Kortum, "In vivo fluorescence hyperspectral imaging of oral neoplasia," in *Proc. SPIE 7169, Advanced Biomedical and Clinical Diagnostic Systems VII*, 2009, p. 71690J.

[123] D. Roblyer, C. Kurachi, V. Stepanek, M. D. Williams, A. K. El-Naggar, J. J. Lee, *et al.*, "Objective Detection and Delineation of Oral Neoplasia Using Autofluorescence Imaging," *Cancer Prevention Research,* vol. 2, pp. 423-431, 2009.

[124] M. Isabelle, K. Rogers, and N. Stone, "Correlation mapping: rapid method for identification of histological features and pathological classification in mid infrared spectroscopic images of lymph nodes," *Journal of Biomedical Optics,* vol. 15, pp. 026030-026030-5, 2010.

[125] A. J. Chaudhari, F. Darvas, J. R. Bading, R. A. Moats, P. S. Conti, D. J. Smith, *et al.*, "Hyperspectral and multispectral bioluminescence optical tomography for small animal imaging," *Phys Med Biol,* vol. 50, pp. 5421-41, 2005.

[126] J. W. Uhr, M. L. Huebschman, E. P. Frenkel, N. L. Lane, R. Ashfaq, H. Liu, *et al.*, "Molecular profiling of individual tumor cells by hyperspectral microscopic imaging," *Translational Research,* vol. 159, pp. 366-375, 2012.

[127] J. Freeman, F. Downs, L. Marcucci, E. N. Lewis, B. Blume, J. Rish, *et al.*, "Multispectral and hyperspectral imaging: Applications for medical and surgical diagnostics," in *Proceedings of the 19th Annual International Conference of the Ieee Engineering in Medicine and Biology Society*. vol. 19, ed, 1997, pp. 700-701.

[128] K. J. Zuzak, R. P. Francis, E. F. Wehner, M. Litorja, J. A. Cadeddu, and E. H. Livingston, "Active DLP hyperspectral illumination: a noninvasive, in vivo, system characterization visualizing tissue oxygenation at near video rates," *Anal Chem*, vol. 83, pp. 7424-30, 2011.

[129] E. Wehner, A. Thapa, E. Livingston, and K. Zuzak, "NIR DLPÃ Â® hyperspectral imaging system for medical applications," in *Proc. SPIE 7932, Emerging Digital Micromirror Device Based Systems and Applications III*, 2011, pp. 793204-793204-9.

[130] K. J. Zuzak, S. C. Naik, G. Alexandrakis, D. Hawkins, K. Behbehani, and E. Livingston, "Intraoperative bile duct visualization using near-infrared hyperspectral video imaging," *The American Journal of Surgery*, vol. 195, pp. 491-497, 2008.

[131] K. Mitra, J. Melvin, S. Chang, K. Park, A. Yilmaz, S. Melvin, *et al.*, "Indocyanine-green-loaded microballoons for biliary imaging in cholecystectomy," *Journal of Biomedical Optics*, vol. 17, p. 116025, 2012.

[132] M. S. Holzer, S. L. Best, N. Jackson, A. Thapa, G. V. Raj, J. A. Cadeddu, *et al.*, "Assessment of Renal Oxygenation During Partial Nephrectomy Using Hyperspectral Imaging," *The Journal of urology*, vol. 186, pp. 400-404, 2011.

[133] E. O. Olweny, S. Faddegon, S. L. Best, N. Jackson, E. F. Wehner, Y. K. Tan, *et al.*, "Renal oxygenation during robotic-assisted laparoscopic partial nephrectomy: characterization using laparoscopic digital light processing hyperspectral imaging," *Journal of Endourology*, vol. 27, pp. 265-269, 2013.

[134] K. J. Zuzak, R. P. Francis, E. F. Wehner, J. Simith, M. Litorja, D. W. Allen, *et al.*, "Hyperspectral imaging utilizing LCTF and DLP technology for surgical and clinical applications," in *Proc. SPIE 7170, Design and Quality for Biomedical Technologies II*, 2009, p. 71700C.

[135] K. J. Zuzak, R. P. Francis, E. F. Wehner, J. Smith, M. Litorja, D. W. Allen, *et al.*, "DLP hyperspectral imaging for surgical and clinical utility," in *Proc. SPIE 7210,*

Emerging Digital Micromirror Device Based Systems and Applications, 2009, p. 721006.

[136] K. J. Zuzak, E. Wehner, S. Rao, M. Litorja, D. W. Allen, M. Singer, *et al.*, "The robustness of DLP hyperspectral imaging for clinical and surgical utility," in *Proc. SPIE 7596, Emerging Digital Micromirror Device Based Systems and Applications II*, 2010, p. 759604.

[137] H. Akbari, Y. Kosugi, K. Kojima, and N. Tanaka, "Hyperspectral imaging and diagnosis of intestinal ischemia," in *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, 2008, pp. 1238-1241.

[138] V.-D. Tuan, "A hyperspectral imaging system for in vivo optical diagnostics," *Engineering in Medicine and Biology Magazine, IEEE,* vol. 23, pp. 40-49, 2004.

[139] P. Zhihong, G. Healey, M. Prasad, and B. Tromberg, "Face recognition in hyperspectral images," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 25, pp. 1552-1560, 2003.

[140] A. Kienle, L. Lilge, M. S. Patterson, R. Hibst, R. Steiner, and B. C. Wilson, "Spatially resolved absolute diffuse reflectance measurements for noninvasive determination of the optical scattering and absorption coefficients of biological tissue," *Applied Optics,* vol. 35, pp. 2304-2314, 1996.

[141] I. Georgakoudi, B. C. Jacobson, J. Van Dam, V. Backman, M. B. Wallace, M. G. Müller, *et al.*, "Fluorescence, reflectance, and light-scattering spectroscopy for evaluating dysplasia in patients with Barrett's esophagus," *Gastroenterology,* vol. 120, pp. 1620-1629, 2001.

[142] I. J. Bigio and J. R. Mourant, "Ultraviolet and visible spectroscopies for tissue diagnostics: fluorescence spectroscopy and elastic-scattering spectroscopy," *Physics in Medicine and Biology,* vol. 42, p. 803, 1997.

[143] R. Doornbos, R. Lang, M. Aalders, F. Cross, and H. Sterenborg, "The determination of in vivo human tissue optical properties and absolute chromophore concentrations using spatially resolved steady-state diffuse reflectance spectroscopy," *Physics in medicine and biology,* vol. 44, p. 967, 1999.

[144] M. G. Müller, T. A. Valdez, I. Georgakoudi, V. Backman, C. Fuentes, S. Kabani, *et al.*, "Spectroscopic detection and evaluation of morphologic and biochemical changes in early human oral carcinoma," *Cancer,* vol. 97, pp. 1681-1692, 2003.

[145] Chung-Chieh Yu, G. O. D. Condon Lau, Jelena Mirkovic, Sasha McGee, Luis Galindo, Alphi Elackattu, *et al.*, "Quantitative spectroscopic imaging for non-invasive early cancer detection," *Optical Society of America,* vol. 16, pp. 16227-16239, 2008.

[146] M. E. Martin, M. Wabuyele, M. Panjehpour, B. Overholt, R. DeNovo, S. Kennel, *et al.*, "An AOTF-based dual-modality hyperspectral imaging system (DMHSI) capable of simultaneous fluorescence and reflectance imaging," *Medical Engineering & Physics,* vol. 28, pp. 149-155, 2006.

[147] A. L. Vahrmeijer, M. Hutteman, J. R. van der Vorst, C. J. van de Velde, and J. V. Frangioni, "Image-guided cancer surgery using near-infrared fluorescence," *Nature Reviews Clinical Oncology,* vol. 10, pp. 507-518, 2013.

[148] A. O. Gerstner, "Early detection in head and neck cancer - current state and future perspectives," *GMS current topics in otorhinolaryngology - head and neck surgery,* vol. 7, 2008.

[149] G. Lu, L. Halig, D. Wang, X. Qin, Z. G. Chen, and B. Fei, "Spectral-spatial classification for noninvasive cancer detection using hyperspectral imaging," *J Biomed Opt,* vol. 19, p. 106004, Oct 2014.

[150] X. Qin, Z. Cong, L. V. Halig, and B. Fei, "Automatic Segmentation of Right Ventricle on Ultrasound Images Using Sparse Matrix Transform and Level Set," in *Proc. SPIE 8669, Medical Imaging 2013: Image Processing*, 2013, p. 86690Q.

[151] X. Qin, Z. Cong, and B. Fei, "Automatic segmentation of right ventricular ultrasound images using sparse matrix transform and a level set," *Physics in Medicine and Biology,* vol. 58, p. 7609, 2013.

[152] N. Renard and S. Bourennane, "Improvement of Target Detection Methods by Multiway Filtering," *Geoscience and Remote Sensing, IEEE Transactions on,* vol. 46, pp. 2407-2417, 2008.

[153] T. Lin and S. Bourennane, "Hyperspectral Image Processing by Jointly Filtering Wavelet Component Tensor," *Geoscience and Remote Sensing, IEEE Transactions on,* vol. 51, pp. 3529-3541, 2013.

[154] N. Renard, S. Bourennane, and J. Blanc-Talon, "Denoising and Dimensionality Reduction Using Multilinear Tools for Hyperspectral Images," *Geoscience and Remote Sensing Letters, IEEE,* vol. 5, pp. 138-142, 2008.

[155] N. Renard and S. Bourennane, "Dimensionality Reduction Based on Tensor Modeling for Classification Methods," *Geoscience and Remote Sensing, IEEE Transactions on,* vol. 47, pp. 1123-1131, 2009.

[156] S. Bourennane, C. Fossati, and A. Cailly, "Improvement of Classification for Hyperspectral Images Based on Tensor Modeling," *Geoscience and Remote Sensing Letters, IEEE,* vol. 7, pp. 801-805, 2010.

[157] Z. Liangpei, Z. Lefei, T. Dacheng, and H. Xin, "Tensor Discriminative Locality Alignment for Hyperspectral Image Spectral-Spatial Feature Extraction," *Geoscience and Remote Sensing, IEEE Transactions on,* vol. 51, pp. 242-256, 2013.

[158] S. Hemissi, I. R. Farah, K. Saheb Ettabaa, and B. Solaiman, "Multi-Spectro-Temporal Analysis of Hyperspectral Imagery Based on 3-D Spectral Modeling and Multilinear Algebra," *Geoscience and Remote Sensing, IEEE Transactions on,* vol. 51, pp. 199-216, 2013.

[159] S. Velasco-Forero and J. Angulo, "Classification of hyperspectral images by tensor modeling and additive morphological decomposition," *Pattern Recognition,* vol. 46, pp. 566-577, 2013.

[160] A. H. Phan and A. Cichocki, "Tensor decompositions for feature extraction and classification of high dimensional datasets," *Nonlinear Theory and Its Applications, IEICE,* vol. 1, pp. 37-68, 2010.

[161] X. Zhang, L. Su, A. Pirani, H. Wu, H. Zhang, D. Shin, *et al.*, "Understanding metastatic SCCHN cells from unique genotypes to phenotypes with the aid of an animal model and DNA microarray analysis," *Clinical & Experimental Metastasis,* vol. 23, pp. 209-222, 2006.

[162] E. Claridge and D. Hidovic-Rowe, "Model based inversion for deriving maps of histological parameters characteristic of cancer from ex-vivo multispectral images of the colon," *IEEE Trans Med Imaging,* vol. 33, pp. 822-35, 2014.

[163] X. Qin and B. Fei, "Measuring myofiber orientations from high-frequency ultrasound images using multiscale decompositions," *Physics in Medicine and Biology,* vol. 59, p. 3907, 2014.

[164] X. Qin, Z. Cong, R. Jiang, M. Shen, M. B. Wagner, P. Kishbom, *et al.*, "Extracting Cardiac Myofiber Orientations from High Frequency Ultrasound Images," in *Proc. SPIE 8675, Medical Imaging 2013: Ultrasonic Imaging, Tomography, and Therapy*, 2013, p. 867507.

[165] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM review,* vol. 51, pp. 455-500, 2009.

[166] A. Cichocki, "Tensor Decompositions: A New Concept in Brain Data Analysis?," *Journal of SICE Control Measurement, and System Integration, special issue; Measurement of Brain Functions and Bio-Signals,* vol. 50, pp. 507-517, 2011.

[167] C.-C. Chang and C.-J. Lin, "LIBSVM: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology (TIST),* vol. 2, p. 27, 2011.

[168] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision,* vol. 1, pp. 321-331, 1988.

[169] X. Qin, S. Wang, and M. Wan, "Improving reliability and accuracy of vibration parameters of vocal folds based on high-speed video and electroglottography," *IEEE Trans Biomed Eng,* vol. 56, pp. 1744-54, 2009.

[170] T. F. Chan, B. Y. Sandberg, and L. A. Vese, "Active Contours without Edges for Vector-Valued Images," *Journal of Visual Communication and Image Representation,* vol. 11, pp. 130-141, 2000.

[171] C. Rodarmel and J. Shan, "Principal component analysis for hyperspectral image classification," *Surveying and Land Information Science,* vol. 62, pp. 115-122, 2002.

[172] G. Lv, G. Yan, and Z. Wang, "Bleeding detection in wireless capsule endoscopy images based on color invariants and spatial pyramids using support vector machines," IEEE Eng Med Biol Soc, 2011.

[173] X. Qin, G. Lu, I. Sechopoulos, and B. Fei, "Breast tissue classification in digital tomosynthesis images based on global gradient minimization and texture features," in *Proc. SPIE 9034, Medical Imaging 2014: Image Processing*, 2014, p. 90341V.

[174] N. Bedard, R. A. Schwarz, A. Hu, V. Bhattar, J. Howe, M. D. Williams, *et al.*, "Multimodal snapshot spectral imaging for oral cancer diagnostics: a pilot study," *Biomedical Optics Express,* vol. 4, pp. 938-949, 2013.

[175] M. M. Stephen, J. L. Jayanthi, N. G. Unni, P. E. Kolady, V. T. Beena, P. Jeemon, *et al.*, "Diagnostic accuracy of diffuse reflectance imaging for early detection of pre-malignant and malignant changes in the oral cavity: a feasibility study," *BMC Cancer,* vol. 13, p. 278, 2013.

[176] A. Bashkatov, E. Genina, V. Kochubey, and V. Tuchin, "Optical properties of human skin, subcutaneous and mucous tissues in the wavelength range from 400 to 2000 nm," *Journal of Physics D: Applied Physics,* vol. 38, p. 2543, 2005.

[177] J. Ferlay, I. Soerjomataram, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, *et al.*, "Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012," *International Journal of Cancer,* vol. 136, pp. E359-E386, 2015.

[178] L. G. T. Morris, A. G. Sikora, S. G. Patel, R. B. Hayes, and I. Ganly, "Second Primary Cancers After an Index Head and Neck Cancer: Subsite-Specific Trends in the Era of Human Papillomavirus–Associated Oropharyngeal Cancer," *Journal of Clinical Oncology,* vol. 29, pp. 739-746, 2011.

[179] D. V. Messadi, "Diagnostic aids for detection of oral precancerous conditions," *Int J Oral Sci,* vol. 5, pp. 59-65, 2013.

[180] A. Rashid and S. Warnakulasuriya, "The use of light-based (optical) detection systems as adjuncts in the detection of oral cancer and oral potentially malignant disorders: a systematic review," *Journal of Oral Pathology & Medicine,* vol. 44, pp. 307-328, 2015.

[181] D. Roblyer, R. Richards-Kortum, K. Sokolov, A. K. El-Naggar, M. D. Williams, C. Kurachi, *et al.*, "Multispectral optical imaging device for in vivo detection of oral neoplasia," *J Biomed Opt*, vol. 13, p. 024019, 2008.

[182] A. Hellebust, K. Rosbach, J. K. Wu, J. Nguyen, A. Gillenwater, N. Vigneswaran, *et al.*, "Vital-dye-enhanced multimodal imaging of neoplastic progression in a mouse model of oral carcinogenesis," *J Biomed Opt*, vol. 18, p. 126017, 2013.

[183] D. Kanojia and M. M. Vaidya, "4-Nitroquinoline-1-oxide induced experimental oral carcinogenesis," *Oral Oncology*, vol. 42, pp. 655-667, 2006.

[184] P. M. Speight, "Update on Oral Epithelial Dysplasia and Progression to Cancer," *Head and Neck Pathology*, vol. 1, pp. 61-66, 2007.

[185] D. Roblyer, C. Kurachi, V. Stepanek, M. D. Williams, A. K. El-Naggar, J. J. Lee, *et al.*, "Objective Detection and Delineation of Oral Neoplasia Using Autofluorescence Imaging," *Cancer prevention research (Philadelphia, Pa.)*, vol. 2, pp. 423-431, 2009.

[186] G. Lu, L. Halig, D. Wang, Z. G. Chen, and B. Fei, "Spectral-Spatial Classification Using Tensor Modeling for Cancer Detection with Hyperspectral Imaging," in *Proc SPIE*, 2014, p. 903413.

[187] D. Roblyer, C. Kurachi, V. Stepanek, R. A. Schwarz, M. D. Williams, A. K. El-Naggar, *et al.*, "Comparison of multispectral wide-field optical imaging modalities to maximize image contrast for objective discrimination of oral neoplasia," *J Biomed Opt*, vol. 15, p. 066017, 2010.

[188] N. Subhash, J. R. Mallia, S. S. Thomas, A. Mathews, P. Sebastian, and J. Madhavan, "Oral cancer detection using diffuse reflectance spectral ratio R540/R575 of oxygenated hemoglobin bands," *J Biomed Opt*, vol. 11, p. 014018, 2006.

[189] G. Lu and B. Fei, "Medical hyperspectral imaging: a review," *Journal of Biomedical Optics*, vol. 19, pp. 010901-010901, 2014.

[190] B. C. Wilson and S. L. Jacques, "Optical reflectance and transmittance of tissues: principles and applications," *IEEE Journal of Quantum Electronics*, vol. 26, pp. 2186-2199, 1990.

[191] L. Barnes, J. Eveson, and D. Sidransky, *Pathology and genetics of head and neck tumours*: Lyon: IARC Press, 2005.

[192] T. D. Gautheir, "Detecting Trends Using Spearman's Rank Correlation Coefficient," *Environmental Forensics,* vol. 2, pp. 359-362, 2001.

[193] M. M. R. Krishnan, V. Venkatraghavan, U. R. Acharya, M. Pal, R. R. Paul, L. C. Min, *et al.*, "Automated oral cancer identification using histopathological images: A hybrid feature extraction paradigm," *Micron,* vol. 43, pp. 352-364, 2012.

[194] H. W. Beumer, K. Vishwanath, L. Puscas, H. R. Afshari, N. Ramanujam, and W. T. Lee, "Detection of squamous cell carcinoma and corresponding biomarkers using optical spectroscopy," *Otolaryngol Head Neck Surg,* vol. 144, pp. 390-4, 2011.

[195] T. A. Winning and G. C. Townsend, "Oral mucosal embryology and histology," *Clinics in Dermatology,* vol. 18, pp. 499-511.

[196] P. M. Speight, P. M. Farthing, and J. E. Bouquot, "The pathology of oral cancer and precancer," *Current Diagnostic Pathology,* vol. 3, pp. 165-176, 1996.

[197] S. Keereweer, H. J. Sterenborg, J. D. Kerrebijn, P. B. Van Driel, R. J. Baatenburg de Jong, and C. W. Lowik, "Image-guided surgery in head and neck cancer: current practice and future directions of optical imaging," *Head Neck,* vol. 34, pp. 120-6, 2012.

[198] R. H. Spiro, O. Guillamondegui, Jr., A. F. Paulino, and A. G. Huvos, "Pattern of invasion and margin assessment in patients with oral tongue cancer," *Head Neck,* vol. 21, pp. 408-13, 1999.

[199] L. Jacobs, "Positive Margins: The Challenge Continues for Breast Surgeons," *Annals of Surgical Oncology,* vol. 15, pp. 1271-1272, 2008.

[200] R. Haque, R. Contreras, M. P. McNicoll, E. C. Eckberg, and D. B. Petitti, "Surgical margins and survival after head and neck cancer surgery," *BMC Ear Nose Throat Disord,* vol. 6, p. 2, 2006.

[201] F. Meric, N. Q. Mirza, G. Vlastos, T. A. Buchholz, H. M. Kuerer, G. V. Babiera, *et al.*, "Positive surgical margins and ipsilateral breast tumor recurrence predict

disease-specific survival after breast-conserving therapy," *Cancer,* vol. 97, pp. 926-33, 2003.

[202] S. Kiyotoki, J. Nishikawa, T. Okamoto, K. Hamabe, M. Saito, A. Goto, *et al.*, "New method for detection of gastric cancer by hyperspectral imaging: a pilot study," *Journal of Biomedical Optics,* vol. 18, pp. 026010-026010, 2013.

[203] S. C. Gebhart, R. C. Thompson, and A. Mahadevan-Jansen, "Liquid-crystal tunable filter spectral imaging for brain tumor demarcation," *Applied Optics,* vol. 46, pp. 1896-1910, 2007.

[204] K. J. Zuzak, S. C. Naik, G. Alexandrakis, D. Hawkins, K. Behbehani, and E. Livingston, "Intraoperative bile duct visualization using near-infrared hyperspectral video imaging," *Am J Surg,* vol. 195, pp. 491-7, 2008.

[205] D. Wang, S. Müller, A. R. M. R. Amin, D. Huang, L. Su, Z. Hu, *et al.*, "The Pivotal Role of Integrin β1 in Metastasis of Head and Neck Squamous Cell Carcinoma," *Clinical Cancer Research,* vol. 18, pp. 4589-4599, 2012.

[206] H. Lange, "Automatic glare removal in reflectance imagery of the uterine cervix," *Proc. SPIE 5747,* vol. 5747, pp. 2183-2192, 2005.

[207] G. Lu, X. Qin, D. Wang, Z. G. Chen, and B. Fei, "Quantitative wavelength analysis and image classification for intraoperative cancer diagnosis with hyperspectral imaging," in *Proc. SPIE* 2015, p. 94151B.

[208] K.-M. Saipullah and D.-H. Kim, "Target detection of hyperspectral images based on their Fourier spectral features," *Optical Engineering,* vol. 51, pp. 111704-1-111704-11, 2012.

[209] R. Pourreza-Shahri, F. Saki, N. Kehtarnavaz, P. LeBoulluec, and H. Liu, "Classification of ex-vivo breast cancer positive margins measured by hyperspectral imaging," in *Image Processing (ICIP), 2013 20th IEEE International Conference on,* 2013, pp. 1408-1412.

[210] H. Peng, L. Fulmi, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 27, pp. 1226-1238, 2005.

[211] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *Systems, Man and Cybernetics, IEEE Transactions on,* vol. 9, pp. 62-66, 1979.

[212] J. N. Kapur, P. K. Sahoo, and A. K. C. Wong, "A new method for gray-level picture thresholding using the entropy of the histogram," *Computer Vision, Graphics, and Image Processing,* vol. 29, pp. 273-285, 1985.

[213] N. Lue, J. W. Kang, C. C. Yu, I. Barman, N. C. Dingari, M. S. Feld, *et al.*, "Portable optical fiber probe-based spectroscopic scanner for rapid cancer diagnosis: a new tool for intraoperative margin assessment," *PLoS One,* vol. 7, p. e30887, 2012.

[214] D. N. Sutton, J. S. Brown, S. N. Rogers, E. D. Vaughan, and J. A. Woolgar, "The prognostic implications of the surgical margin in oral squamous cell carcinoma," *International Journal of Oral and Maxillofacial Surgery,* vol. 32, pp. 30-34, 2003.

[215] M. L. Hinni, A. Ferlito, M. S. Brandwein-Gensler, R. P. Takes, C. E. Silver, W. H. Westra, *et al.*, "Surgical margins in head and neck cancer: a contemporary review," *Head Neck,* vol. 35, pp. 1362-70, 2013.

[216] L. J. DiNardo, J. Lin, L. S. Karageorge, and C. N. Powers, "Accuracy, utility, and cost of frozen section margins in head and neck cancer surgery," *Laryngoscope,* vol. 110, pp. 1773-6, 2000.

[217] E. Du, T. J. Ow, Y.-T. Lo, A. Gersten, B. A. Schiff, A. B. Tassler, *et al.*, "Refining the utility and role of Frozen section in head and neck squamous cell carcinoma resection," *The Laryngoscope,* pp. 1768-1775, 2016.

[218] G. Lu, X. Qin, D. Wang, M. S., Z. H., A. Chen, *et al.*, "Hyperspectral Imaging of Neoplastic Progression in a Mouse Model of Oral Carcinogenesis," in *Proc SPIE* 2016, p. 978812.

[219] F. Baowei, H. Akbari, and L. V. Halig, "Hyperspectral imaging and spectral-spatial classification for cancer detection," *Biomedical Engineering and Informatics (BMEI), 2012 5th International Conference on,* pp. 62-64, 2012.

[220] G. Lu, D. Wang, X. Qin, L. Halig, S. Muller, H. Zhang, *et al.*, "Framework for hyperspectral image processing and quantification for cancer detection during animal tumor surgery," *J Biomed Opt,* vol. 20, p. 126012, 2015.

[221] Z. Luo, M. N. Loja, D. G. Farwell, Q. C. Luu, P. J. Donald, D. Amott, *et al.*, "Widefield Optical Imaging of Changes in Uptake of Glucose and Tissue Extracellular pH in Head and Neck Cancer," *Cancer Prevention Research,* vol. 7, pp. 1035-1044, 2014.

[222] P. Vila, C. Park, M. Pierce, G. Goldstein, L. Levy, V. Gurudutt, *et al.*, "Discrimination of Benign and Neoplastic Mucosa with a High-Resolution Microendoscope (HRME) in Head and Neck Cancer," *Annals of Surgical Oncology,* vol. 19, pp. 3534-3539, 2012.

[223] G. Lu, L. Halig, D. Wang, Z. G. Chen, and B. Fei, "Hyperspectral imaging for cancer surgical margin delineation: registration of hyperspectral and histological images," in *Proc. SPIE 9036, Medical Imaging 2014: Image-Guided Procedures, Robotic Interventions, and Modeling*, 2014, p. 90360S.

VITA

Guolan Lu was born in Suizhou, Hubei Province in China. She received her bachelor's degree in Measurement Control Technology and Instrumentation from Xidian University in Xi'an, China, in 2009. She earned a master's degree in Precision Instrument and Machinery from Shanghai Jiao Tongue University and a master's degree in Electrical and Computer Engineering from Georgia Institute of Technology in the spring of 2012. Starting in the fall of 2012, she continued to pursue her doctorate degree in the Wallace H. Coulter Department of Biomedical Engineering at Georgia Tech and Emory University. Since then, she has been a graduate research assistant in the Quantitative BioImaging Laboratory (QBIL) under the guidance of Professor Baowei Fei and her doctorate research has focused on hyperspectral imaging and machine learning for early cancer detection and image-guided surgery.